AN INVESTIGATION OF THE RELATIVE EFFECTIVENESS OF THE PERSONALIZED SYSTEM OF INSTRUCTION AT THE NAVAL POSTGRADUATE SCHOOL

Patrick Alan Toffler

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

AN INVESTIGATION OF THE RELATIVE
EFFECTIVENESS OF THE
PERSONALIZED SYSTEM OF INSTRUCTION
AT THE
NAVAL POSTGRADUATE SCHOOL

by

Patrick Alan Toffler

September 1975

Thesis Advisor: M. D. Weir

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) An Investigation of the Relative Effectiveness of the Personalized System of Instruction at the Naval Postgraduate School		5. TYPE OF REPORT & PERIOD COVERED Master's Thesis; September 1975
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Patrick Alan Toffler		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		12. REPORT DATE September 1975
		13. NUMBER OF PAGES 113
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Personalized System of Instruction (PSI)
Education and Training

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

An extensive and scientifically controlled experiment was designed and conducted over a three month period. The analysis of the results are accomplished with modern, advanced statistical procedures. The effort leads to the conclusion that the Personalized System of Instruction (PSI) is demonstrably superior to the conventional lecture/recitation (CLR) technique for teaching graduate level students (at NPS in a certain class of subjects). This paper describes in detail the nature of the experiment, the analysis and the benefits

DD FORM 1 JAN 73 1473 (Page 1) EDITION OF 1 NOV 65 IS OBSOLETE S/N 0102-014-6601 |

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) An Investigation of the Relative Effectiveness of the Personalized System of Instruction at the Naval Postgraduate School		5. TYPE OF REPORT & PERIOD COVERED Master's Thesis; September 1975
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Patrick Alan Toffler		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		12. REPORT DATE September 1975
		13. NUMBER OF PAGES 113
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Personalized System of Instruction (PSI)
Education and Training

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

An extensive and scientifically controlled experiment was designed and conducted over a three month period. The analysis of the results are accomplished with modern, advanced statistical procedures. The effort leads to the conclusion that the Personalized System of Instruction (PSI) is demonstrably superior to the conventional lecture/recitation (CLR) technique for teaching graduate level students (at NPS in a certain class of subjects). This paper describes in detail the nature of the experiment, the analysis and the benefits

ABSTRACT

An extensive and scientifically controlled experiment was designed and conducted over a three month period. The analysis of the results are accomplished with modern, advanced statistical procedures. The effort leads to the conclusion that the Personalized System of Instruction (PSI) is demonstrably superior to the conventional lecture/recitation (CLR) technique for teaching graduate level students (at NPS in a certain class of subjects). This paper describes in detail the nature of the experiment, the analysis and the benefits to be derived through utilization of PSI. The findings of this experiment are directly applicable to the costly and imperative educational and training missions conducted by the Department of Defense (DOD).

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

The author is indebted to many helpful people, without whose guidance and assistance this research effort would not have produced the potentially useful results which it discusses. Among those to whom he is especially grateful are included: Professor Maurice D. Weir, who wrote the course materials for the PSI Class; LT Gerald G. Brown who assisted in the design of the experiment and patiently guided the analysis of the data; Professor Alan Shorb, who taught the CLR group and managed the PSI group, his efforts to develop this analyst as an effective tutor are much appreciated; and Professor Gary Poock without whose help the proper construction of the Analysis of Variance would have taken much longer than it did. Special thanks is offered to the tireless staff of the W. R. Church Computer Center, Naval Postgraduate School, for their patience and technical assistance. The invaluable contribution of Gerald P. Learmonth resulted in the ability to employ sophisticated data analysis techniques, only recently developed, to extract meaningful information from the data base.

No thesis on this subject could have been written if Dr. Ben A. Green, Center for PSI, Georgetown University, had not worked so diligently to bring PSI from the conceptual stage into the environment where it could mature and begin to realize its potential. His pioneering efforts and courage to stand behind an innovative idea are respectfully acknowledged.

The author bears full responsibility for the manner in which the research and analysis was conducted and for the conclusions and recommendations advanced herein.

FOREWORD

The educational effort supported by the Department of Defense (DOD) is very great. Every year billions of dollars are expended to produce graduates from hundreds of schools who are hopefully trained and ready to perform one of many specialty functions. It is of great interest, to those involved in this process, when the chance to improve the quality and efficiency of the educational process presents itself.

An innovative instructional methodology, called the Personalized System of Instruction (PSI), burst upon the academic scene in the mid-60's. It promised to markedly enhance the benefits which graduates of courses could derive from their educational experience.

This paper discusses an analytic investigation of the effects and benefits which can accrue to students learning via PSI. It is sincerely hoped that the findings which it details will assist in the evaluation of this new technique for its applicability to the Department of Defense.

I. INTRODUCTION

A. LEAD-IN

Ever since man began teaching others he has wondered about and sought methods for improving the efficiency and effectiveness of the educational process. In spite of these efforts education has sustained little systematic change: teaching has remained essentially an art. Controversies have arisen concerning teaching methodology and the proportion of responsibility that rests with the instructor and with his students. These moot issues have remained a matter for subjective analysis over the years because the field of education is only beginning to emerge as a scientific discipline. Recently discovered knowledge of learning principles derived from experimental analysis of human behavior has had an important and dramatic affect on the development of a scientifically based technology of teaching [Skinner, 1968]. One effort in this regard is described by Keller [1966], and further research on college teaching corroborates his hypothesis that the application of behavioral principles in the instructional environment does yield significant results in higher education [Ferster, 1968; McMichael and Corey, 1969; Sheppard and MacDernott, 1970; Alba and Pennypacker, 1971, 1972; Kulik, 1975].

The need for continuing study and development of more effective and efficient teaching techniques is manifest, not only because of the

growing number of students requiring an education, but also because of the complexity, diversity and quantity of the material they must learn. In addition budgetary considerations are becoming more constrictive as the need for education competes, for funds, with other requirements of great urgency or high priority. Thus, the problem of maximizing teaching effectiveness most efficiently (i.e., at lowest cost) is a critical one.

Therefore, if the new findings concerning the nature of the learning process hold the promise of a true advance in education then they should be fully explored. The research effort to be described in this thesis was conducted to aid in this regard and, particularly, to address the question of whether a technique termed the Personalized System of Instruction (PSI) [Keller, Sherman, 1974] affords specific educational advantages to the Department of Defense (DOD).

B. SITUATION

The Naval Postgraduate School (NPS) trains military officers in various scientific disciplines at the graduate university level. The curricula are demanding and the students matriculate with widely varying academic backgrounds. Frequently it has been many years since they were exposed to college level academic work. The challenge of educating these men as quickly and as effectively as possible is a matter of great interest in an era of budgetary restrictions and increasing technological demands placed upon officers during the conduct of

their operational duties. The question of applying PSI as an instructional technique within the NPS program was raised by the administration and a few courses were designed, written and administered via this mode. This thesis analyzes the design, conduct and outcome of one such experiment (a course in linear algebra). This experiment and its findings are of importance for several reasons. First, the research effort was conducted from the start for the purpose of addressing the question of PSI's relative effectiveness and its potential. The specific output measures employed were carefully selected to provide meaningful information that is reliable and valid. Second, the experiment was meticulously controlled. All relevant, influential variables were held constant save for the form of instruction which was left free to affect the students, independent of contaminating influences. Third, the experiment was a success in that the results were obtained without confounding, missing data points or contamination. The output provides usable information from which decision makers can draw useful conclusions concerning the expected benefits to be derived from the employment of PSI as a teaching mode in DOD educational institutions of many varieties.

It is emphasized that previous experiments presented in the literature, in spite of their praise of PSI, did not address the issues at the graduate level of education with military officers of high calibre as subjects. Further, the great care taken in this experiment to insure the existence of matched groups lends significant credence to the

impressive statistical results observed across the output measures of this experiment. It is genuinely hoped that this research effort will further and speed the application of PSI by the military in its vast and expensive efforts to educate and train its personnel for the conduct of their complex tasks.

II. BACKGROUND

A. CONVENTIONAL LECTURE-RECITATION (CLR)

The academic year at NPS is divided into four (quarter) segments. During each quarter a typical student is exposed to several courses (usually three to five) with credit hours ranging from twelve to twenty-five hours. In addition, students in Masters or Doctoral level programs work on their thesis or dissertations. In a course taught via the conventional (CLR) mode the class meets at designated hours during the week for lectures or laboratory work. Assignments, tests, grading and course content, within specified limits, are the responsibility of the professor. The class moves through the course at a rate determined by the professor and students are afforded the opportunity to meet with him for additional instruction if they desire it. The number of students in a classroom is typically fifteen and sometimes is as high as thirty-five. The amount of "outside class-room work" performed by a student is a function of many variables (e.g., the student's personal goals, his innate academic abilities, his prior exposure to course material, and so forth). This instructional technique is typical of that being practiced in most colleges and universities throughout the United States.

B. THE PERSONALIZED SYSTEM OF INSTRUCTION (PSI)

Consider, for the present, that PSI is a particular mode of instruction that differs from CLR in format but not in purpose. That is, PSI and CLR have as their primary objective the goal of insuring that students learn as much as possible about the subject matter at hand (i. e., theory, mechanics, subtleties, construct, etc.) in a reasonable amount of time. Given this supposition it is important to clarify the true nature of PSI and to illustrate the manner in which it differs from CLR.

In March 1963 four psychologists, advocates of the REINFORCEMENT THEORY in learning and adherents of many of the Skinnerian Principles of Human Behavior [Friedman, et al 1975] conceived a method of instruction that departed from traditional methods which they believed were outdated [Keller, Sherman, 1974]. Although no single feature of their design was original, the combination of the various features was to produce perhaps "one of the most exciting and radical (educational innovations) ever introduced at the university level [Ibid]. As the method developed, its skeletal form embodied certain fundamental features: the breakdown of the course into small learning "units", self-pacing, repeated testing, and mastery learning before exposure to sequential unit material. A PSI course might contain lectures, demonstrations, discussions, laboratory/workshop hours, and homework. However, the lecture and demonstrations were to be infrequent and designed primarily to provide motivation for the student rather than to act as sources of essential information. These lectures

were to be interesting, informative and memorable -- even entertaining. They were producible at suitable intervals and available to students who had progressed to a point that optimized their ability to appreciate their content. Attendance at these enrichment activities was optional and their material was not intended for testing.

As PSI matured other salient features became manifest. The role of the professor evolved to that of an educational engineer, a contingency manager, whose principal function was to facilitate the learning process. The presence and function of proctors (or tutors) to assist the professor in grading and answering individual student questions became paramount. The proctors were to interface directly with the students clarifying complex issues from the written materials or the text, diagnosing intermediate student performance, and prescribing remedial learning activities designed by the professor. Details of these various facets of PSI are explained in Part III.

Beneath all of the technical characteristics of PSI rest the philosophical concepts of behavioral psychology that motivate the design of the new format. The basic strategy inherent in PSI is one that attempts to deal with individual differences in learners: the ability of the learner to comprehend the nature of his task, the manner and order of presentation to optimally tune to the reception frequencies of the learner, the amount of time required by the learner to <u>master</u> the given task (subject to the limitations imposed by the length of the academic segments of the institution). Frequent diagnostic-progress tests provide detailed

feedback to the professor, the tutors and the students: these tests reinforce the learning experience and they reveal particular points of difficulty (which are accompanied by a very specific prescription of what to do about them). No penalty is extracted for non-mastery (save the time required of the student to study harder and then re-attempt to demonstrate a superior degree of proficiency). A final exam could be administered when the student was ready and his grade for the course would be a judicious blending of his previous work and his performance on the final exam. A student who performed poorly on his final might even be afforded an opportunity to take it again.

The rationales upon which these concepts were developed are that the learning of complex materials depends upon the prior mastery of simpler materials [Gangne, 1962, 1965; Moore et al 1973] and that all students do not learn at the same rate [Bloom, 1968]. It was the feeling of the early adherents of PSI that CLR was sub-optimal because it did not ensure the acquisition of simpler concepts and skills before it demanded the absorption of the more complex. Further, by its very nature, CLR pushed (or held) all students through the course material at the same rate.

C. THE VIABILITY OF PSI

Since its birth in 1963 PSI has been tested and adopted at many institutions (e.g., Electricity and Magnetism, Michigan State University; Physics, Bucknell University; Mathematics, New Mexico State University; and others) and has been abandoned only infrequently (e.g., Physics,

Massachusetts Institute of Technology (MIT) [Kulik, 1975, Friedman et al, 1975; Green, 1971]. However, in no two cases was the testing or implementation of PSI exactly the same. Each institution, each professor, and each course differed enough to give the various experiments a character of their own. This fact does not mitigate the desire nor the capacity to examine previous efforts to help establish the benefits of PSI or reveal its weaknesses.

It is important to remember only that any PSI course is built around certain characteristics which distinguish it from CLR. Keller has summarized these aspects as follows:

> 1. The self pace feature, which permits a student to progress through the course at a speed commensurate with his ability and other demands upon his time.
>
> 2. The mastery requirement for advancement to new material, which ensures that students demonstrate understanding of present concepts prior to the study of more complex ideas.
>
> 3. The use of lectures and demonstrations as motivators, rather than sources of critical information.
>
> 4. The stress upon the written word and the "hands-on-the material" concept which ensures that students practice the implementation of the principles they are exposed to.
>
> 5. The use of tutors, who interface with the students in a variety of ways and provide the course with personal/social psychological characteristics.
>
> [Koen, 1974]

Bearing in mind that only courses that possess all of these basic characteristics are considered PSI, it is appropriate to consider some statistics which reveal the demonstrated results achieved with PSI. The task is not as straightforward as it sounds. How does one show that a teaching method works? Stan Ericksen claims that teachers live in a "criterion-free" environment [Kulik, 1975]. If allowance is made for the obvious hyperbole, it is easy to concur with this sentiment. At least it is apparent that, at the present time, there is no ultimate standard against which the efficiency and effectiveness of a teaching process can be measured. Among the useful indicators are: 1) end-of-course performance; 2) retention; 3) transfer; 4) facilitation; and 5) student attitudes. The vast majority of the literature uses one or more of these measures of effectiveness (MOEs) to report the objective results of PSI experiments.

Some reports include a sixth measure: cost-effectiveness - which however, cannot be discussed in a vacuum (i.e., without reference to at least one of the other measures). In order to argue the case for the cost-effectiveness of one alternative vice another, it is necessary to measure the two against some other performance criterion, ascertain what is <u>feasible</u> in terms of resources, and then ask which costs less in order to achieve its demonstrated results. In this paper the cost of PSI is not examined directly. However, no cost-effectiveness analysis could be properly conducted prior to the discoveries of this research, and, without the results discussed here none would be warranted.

A discussion of each of the measures of effectiveness referred to or employed in this study follows:

END OF COURSE PERFORMANCE: Kulik [1975] engaged in considerable independent research and after examining over 400 papers on PSI instructional innovations he located 31 comparisons of final exam scores from PSI and CLR courses. In 30 of these cases final exam (FE) performance by PSI groups was superior to that of CLR classes. In 25 of these cases the difference was statistically reliable ($p = .10$). This means that the probability that the observed differences in group performance were the result of chance alone, given the hypothesis that the two groups were equivalent, was: (.10). Thus the hypothesis of equality of group means could be rejected at the (.90) confidence level. These results are not from isolated studies but represent findings across a broad range of disciplines, course levels and academic institutions [Kulik, 1975].

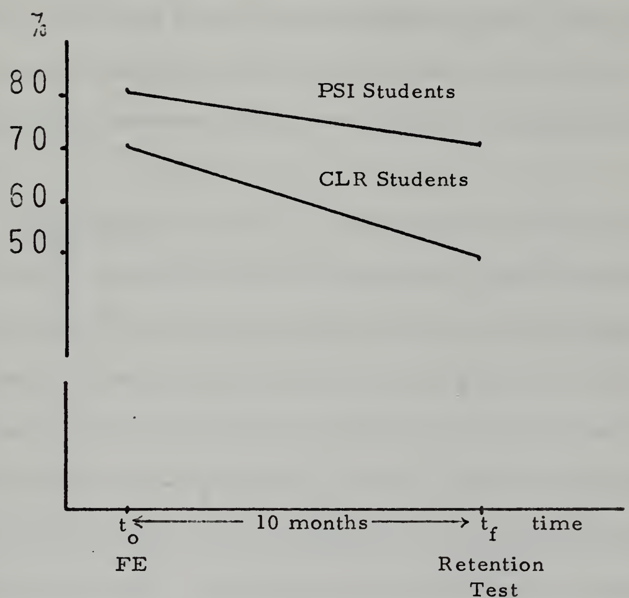
The validity of the FE as a MOE can be questioned. The FE may measure, to some extent, the "amount" that the students have learned but it also reflects the amount that students "cram" and their ability as "test-takers". The scores from the FE can hardly be considered interval data measures of the benefit any student derived from his exposure to the course material, and it is well known that no FE taxes a student's knowledge on all areas with equal reliability or validity. Nonetheless the pervasiveness of the FE in educational research is indisputable.

In the past, all great conflicts over teaching innovations have been fought on the battlefield of the FE. The results have been: stalemate. (For example, a few years ago there were over 80 studies published which compared CLR with the "seminar/discussion" method of instruction. In 51% of these the CLR mode was found to be statistically superior and in the remaining 49% the seminar technique produced significantly better results.) Comparison of other teaching methods using FE as an MOE have ended in the same standoff [Kulik, 1975]. Thus, while the FE is not offered as the "last-word" indicant of performance, in this case, comparison of FE scores shows the PSI groups have obtained a clear-cut margin of victory over their CLR counterparts.

RETENTION: To enrich the analysis of experimental results some investigators have looked beyond the FE. One very careful study at C. W. Post College of Long Island University, New York [Corey, McMichael, 1974], showed that the PSI group retained more information for longer periods and that retention differences between the two groups grew with time (see Figure 1).

In five other studies of retention, where the interval from FE to retention test varied from 3 weeks to 15 months, the PSI groups performed statistically better than students enrolled in CLR classes [Anderson and Artman, 1972; Austin and Gilbert, 1973; Breland and Smith, 1974; Calhoun, 1973; Nazzaro, Todorov, and Nazzari, 1972].

Figure 1. PSI vs CLR Retention Comparison



This figure depicts how PSI students tend to perform better on end of course final exams and how they demonstrate greater retention of learned material over time.

Interpretation of these studies shows that PSI groups seem to be learning information in a manner that facilitates recall over a period of time. Moreover, PSI seems to promote meaningful learning rather than simply the storage of test-related facts in short-term memory. Evidently FE comparisons alone, underestimate the magnitude of PSI effects.

TRANSFER: This MOE is more difficult to measure than its predecessors. It involves the determination of whether information learned via the PSI mode can be applied more readily to related subjects and to practical problems than can information gained in a CLR environment. Four studies which investigated this phenomenon reveal that PSI produces superior transfer effects. Anderson and Artman [1972] found that students who took Physics I via PSI averaged about one letter grade higher than CLR students (from Physics I) when both groups took Physics II via CLR. Students at Michigan State University who took Engineering Statics in a PSI format outperformed their peers from CLR in the follow-on course on Civil Engineering Structures which was administered via CLR [Lubkin, 1974]. Engineering students at the University of Tennessee taking Fluid Mechanics in a PSI mode received better grades than their CLR counterparts when both groups later took Hydraulics via CLR [Weissberg, 1974]. Very recent studies from the University of California, San Diego and from the University of Maryland further confirm the existence of this PSI transfer phenomenon [McMichael, 1975]. Various theories attempting to explain this

phenomenon abound. However, until more is known about how the human brain actually functions, these explanations must remain speculative. Nevertheless, PSI was developed from a knowledge of accepted psychological principles of human learning behavior and evidence is growing in favor of the view that PSI fosters a type of learning that goes beyond that which is characteristically gained from CLR.

FACILITATION: The question of how well PSI students perform in their other (concurrent) courses is an important one. PSI is often charged with requiring an inordinate amount of its students' time and with encouraging them to procrastinate.

A student survey conducted at Ohio State University's College of Medicine addressed this issue [Schimpfhauser, Horrocks, Richardson, Alben, Schumm, and Sprecker, 1974]. All of the medical students considered were enrolled in the same three courses (biochemistry, anatomy and physiology). A group of these students took biochemistry via PSI. All other courses were taken conventionally. Not surprising (to the researchers), the PSI group did better in biochemistry. Somewhat phenomental however, was the result that the PSI students did better than the other students in all their courses. This finding is quite significant since the PSI group was randomly selected and shown to be statistically equivalent to the CLR group over such relevant variables as Grade Point Average (GPA) and Medical College Aptitude Test (MCAT) scores. Most importantly this finding shows that if PSI does make greater demands

upon student study time, it does not do so at the expense of performance in other courses: on the contrary, PSI, with its requirement for mastery, its frequent diagnostic tests, and its unique student-tutor relationship, may foster efficient study habits and imbue in the student a desire to achieve higher standards of performance. This desire to do superior work seems to carry over into all student (academic) activities. Additional data collected by Born and Davis [1974] support that point of view. They found that students do spend more time working on a PSI course than CLR students but that the PSI students spent this time in markedly different ways. Most notable was the absence of time spent listening to lectures and trying (simultaneously) to take meaningful notes. Such activity is largely self-defeating in that the conduct of one action detracts from one's ability to do the other. PSI effectiveness is apparently not derived from the amount of time spent studying but rather in the manner in which this time is spent.

ATTITUDES: Since PSI was conceived in the minds of behavioral psychologists it is not surprising that student perceptions of the course, their sense of personal accomplishment, and their preferences for further learning via PSI was to be an important consideration. Every human experience carries with it some learning about the nature of oneself. Certainly this is true with educational experiences. Every course that the student takes contributes to his self-awareness and his concept of education. This conceptual maturation process is called

attitude growth. Some investigators [Silberman, Parker, 1974] report that PSI contributes to the development of positive attitudes, but objective analysis is difficult and often incomplete. Many researchers have collected open-ended evaluations from their students, and have classified their comments as "positive" or "negative." Judging from published documents the percentage of students commenting unfavorably on the PSI format is very small [approximately 4%, Gallup, 1970; Green, 1970]. When students are asked directly whether they prefer PSI to CLR the results are typically in favor of PSI. At the College of Engineering, University of Texas at Austin, about 70% of the students in a PSI course preferred that mode to CLR: 20% considered the two modes equivalent [Stice, 1971]. Still other reports compare student responses on end of course evaluations. In almost all cases PSI appears highly attractive to students.

The results of these comparisons, across the output measures discussed thus far, were presented by Kulik [1965] and are summarized in Table 1.

Some critics of PSI have attempted to explain away these results by claiming that the HAWTHORNE EFFECT (i. e., an enthusiastic response to something novel) is in operation. However, contrary to this hypothesis, there seems to be a slight increase in the effectiveness of PSI as it matures [Kulik, Kulik, and Miholland, 1974; Roth, 1973], so a sort of anti-HAWTHORNE EFFECT may actually be at work.

Table 1. Comparisons of PSI and CLR Groups

	Favor PSI		Favor CLR		TOTAL
	Statistical Sig (p = .10)	Non-Sig	Statistical Sig (p = .10)	Non-Sig	
FE	25	5	0	1	31
Retention	6	0	0	0	6
Transfer	4	1	0	0	5
Facilitation	1	0	0	0	1
Attitudes	6	1	1	0	8

This table reveals the results of comparisons between PSI and CLR groups taking various courses together across the nation at many different schools. Note: in thirty-one experiments where final exam (FE) scores were used as an output measure that thirty differences in mean score were in favor of the PSI students. Of these thirty differences, twenty-five were so large as to lead one to reject the hypothesis of equality of performance at the 90% confidence level. Only one of the experiments produced results favoring CLR, in the FE category, and that difference was too small to be considered statistically significant at the same confidence level.

There are other arguments that charge PSI materials with "teaching the test." These charges appear weightless when one considers the aforementioned results concerning retention, transfer, and facilitation. Still others submit that the observed results are true but not of any real consequence. Consider an average student, who may take his physics course, for example, via CLR or PSI. If he takes a typical lecture course, his achievement in Physics will place him at the 50th percentile on a standardized test. He is an average student taking an average course. If he takes the same course via PSI, he will achieve at the 75th percentile in the same standardized test. Additionally, he will retain what he learned for a longer period, will find that he can transfer the information to related courses with greater ease, will learn a great deal about efficient study habits and test taking, thus increasing the likelihood that his grades will rise in other courses. This increment - the PSI effect - is what the technique has to offer the individual student. If this effect for one is multiplied by one-hundred, the number of students a typical professor may have in a quarter, then some conception is generated regarding what PSI has to offer the instructor, the academic institution and the taxpayers who support the educational effort.

For most students, most professors, and most decision makers these effects will not seem small.

At this point is is appropriate to consider what PSI really does to the instructor's position. Obviously an educational innovation as radical as PSI would be expected to alter substantially his traditional

role. The question is actually whether the new role is for the better.

Instructors in PSI courses find that they cease being broadcasters of information and that they are no longer "spotlight performers." Instead they become managers of a system and they facilitate learning in others. This point about the systematic character of PSI is noteworthy. Certainly a truly superior CLR course can be constructed and conducted by an outstanding professor. Such a man could duplicate his performance many times (assuming no personal distractions affected his ability to concentrate and prepare himself consistently) but because his course is not systematized it is doubtful that another man could duplicate his efforts with the same course. Not so with PSI. Once the course has been written and validated the professor becomes a system orchestrator. He guides his tutors and develops these subordinates as teachers. He prepares his occasional lectures/demonstrations/workshops and, most important, he can devote himself to helping individual students with particularly significant difficulties. Further, the professor must watch the "dials" in the process: at what rate are students progressing, which students are in trouble and need his special guidance, what logistics problems exist and how can they be mitigated. Thus the competent professor finds that:

a. He knows where each student stands at all times throughout his course.
b. He can give individual attention to students who need him.
c. His students are really mastering his course.
d. He can spend more time preparing his occasional lectures.

e. He can take responsibility for developing his tutors as effective teachers.
f. His administration is pleased that he can teach many students effectively, at low cost.

[Green, 1971].

The professor teaching a PSI course for the first time might feel strange as an "educational engineer," but if he feels his role has diminished because he is not lecturing from his pulpit, on a daily basis, he has perhaps forgotten the central issue in education: the advancement of the student in the discipline of interest.

III. THE EXPERIMENT AT NPS

A. INTRODUCTION

Due to the many indicants of the benefits to be derived from PSI and its sound theoretical basis [Kingsley, 1971] a few courses were recently developed at NPS for presentation both on and off campus. Considerable interest in the merit of these innovations was generated within the faculty, administrative offices, and the student body. Individuals from each of these groups formed opinions based upon varying degrees of exposure to the PSI technique and much subjective analysis was conducted. The whole issue is of considerable interest to the DOD in general. DOD has indicated that NPS should prepare itself for the eventuality that the length of time that students stay on campus for degree completion might be cut (by up to 15%). The rub is that the quality of the education students receive should not be diminished. The requirement to reduce the length of a student's stay on campus is generated by the desire to cut costs. Hundreds of millions of dollars are allocated annually by DOD for training and education. This allocation is warranted as a result of the real need for technically proficient personnel at all levels of activity, including management. However, opportunities to reduce cost levels while maintaining or even improving the quality of the educational process are actively sought on a perennial basis.

Beginning in Quarter II, FY 75 a course written by Professor Maurice D. Weir of the Department of Mathematics was offered to selected students from several disciplines. The course was presented in the PSI mode and adhered to all of the principles of the Keller Plan [Keller, Sherman, 1974]. After its second presentation in Quarter III, FY 75 there was considerable feeling that it was successful. Students were learning well, as evidenced by their scores on their FE as compared with those of students in CLR classes who took the same exam. Additionally, the PSI students expressed appreciation for their exposure to the new mode of instruction and seemed to genuinely like the course. However, many hard questions could only be answered by subjective opinion. Most of these questions had been addressed by experimenters working with other courses at different institutions (as indicated previously), but it was considered appropriate to answer these questions in the unique academic environment at NPS. The major reasons were that the students at NPS were not characteristic of students at most civilian institutions, that the program of instruction and the work load at NPS were considerably more stringent, and because little of the previous research was conducted with graduate students as subjects.

In order to gain an analytic and objective understanding of the effects of PSI at NPS a carefully controlled experiment was designed around Professor Weir's course. This course entitled: "Computational Linear/Matrix Algebra (MA 2045)" provided the basis for the experiment

and the mode in which it was taught constituted the stimulus. The performance of the students, as measured across specific output variables, constituted the response to, or the effect of the stimulus.

In this chapter the design, conduct, results and analysis of that experiment are discussed.

B. THE STUDENTS, THE PROFESSOR, AND THE TUTORS

All of the subjects in this experiment were students in their first quarter at NPS. Their basic disciplines were the technical sciences (e.g., electrical engineering; computer science) and they were working towards the degree Master of Science. All of them were required to take MA 2045 and their selection as members of the PSI group or the CLR group was determined by the administrative offices of NPS (scheduling criteria were the only factors affecting the placement of students in one group or the other). As a result twelve U.S. students and four International students were enrolled in the PSI group. Sixteen U.S. and five International students were scheduled to take the course via CLR. The only assumption concerning these groups was that they were homogeneous or matched (i.e., that across all relevant characteristics the two groups constituted random, independent samples drawn from the same population). It was felt that this was a valid assumption since all students were military officers with fundamentally similar academic and professional experiences. They aspired to similar goals

and had demonstrated personal competence within their career fields prior to selection for attendance at NPS. Hopefully, individual differences between students would balance out within the groups so that the group characteristic centroids would be co-located (in a statistical sense). By this it is acknowledged that the characteristic vector which would completely describe something as complex as a group of students would be multi-dimensional. Since each group consists of several equally complex individuals, the student characteristic vectors, if plotted in the appropriate space, would form a group of points. The center (centroid) of each group should be in approximately the same position for both groups if in fact they are matched.

In order to support the validity of this assumption it was decided, a priori, not to incorporate the performance of the International students in the results of the experiment. This was done because the level of language proficiency varied greatly among these men. To have considered their performance without accounting for language fluency would have introduced an uncontrolled variable in the experiment which would, perhaps, have biased the results.* In addition several measures of U.S. student attributes were obtained and objectively compared to determine the relative validity of the hypothesis of group equality.

---

*The author analyzed the performance of the international students to satisfy his own curiosity. If he had incorporated their performance in his overall analysis the result would have been to even further support the conclusion of increased effectiveness of PSI over CLR.

C. ANALYSIS OF STUDENT CHARACTERISTICS

Six measures of student attributes were obtained prior to any instruction. They were as follows:

a) Performance on a pre-instructional linear/matrix algebra entrance examination;

b) two scores attained on a standardized test of mental maturity [Sullivan, Clark, Tiegs, 1963];

c) two scores attained from a standardized inventory of personality traits [Thorpe, Clark, Tiegs, 1953]; and

d) a numerical representation of the degree of exposure to college level mathematics instruction prior to matriculation at NPS.

The pre-instructional inventory examination was written by Professor Weir: it tested the ability of the students to perform basic manipulations with vectors and matrices. It was administered to determine if the ability of individual students, in one group or the other, was sufficiently developed to permit them to perform in a superior manner in MA 2045 regardless of the form of instruction to which they were exposed. The test of mental maturity [Sullivan et al, 1963] was published by CTB/McGraw-Hill. It was designed to provide information about the functional human abilities that are basic to learning, problem-solving and responding to new situations. The scores produced by this test are similar to those traditionally known as intelligence quotients (I.Q.). The personality inventory [Thorpe, et al 1953], also published by CTB/McGraw-Hill, was administered to help identify the status of

factors in student personality and social adjustment usually designated as "intangibles" (e.g., individual motivation, self-confidence, self-awareness, etc.). These factors traditionally defy appraisal or diagnosis by means of ordinary ability and achievement tests. Both tests published by CTB/McGraw-Hill had been checked for reliability and had been validated over several years. National norms and quantiles were available and it was felt that these tests would provide a satisfactory measure of the variables they were designed to examine. The numerical measure of prior exposure to mathematical subjects was obtained using an algorithm that gave points for the number of, as well as the types of, mathematics courses a student had taken. Then it discounted these point totals by a scalar that was sensitive to the number of years since the courses were taken. (The data obtained from these tests are available to the reader in Appendix B, Sequence 1).

The results of this pre-instructional testing were subsequently analyzed according to several statistical tests of varying power (depending upon the nature of the various assumptions demanded by the different tests). First the data was depicted in histogram form. These graphs reveal that the data, over the variables taken separately, distributes itself in a similar fashion over the groups. However, there is no indication that the nature of this distribution is known. Therefore, the null hypothesis (Ho): that the groups are independent random samples (of size: $m_{psi} = 12$; $n_{clr} = 16$) drawn from the same population, was

tested first according to the non-parametric Kolomogorov-Smirnov (K-S) two-sample criterion, with a computer program that determined the exact probability of the occurrence of the observed difference in group means (or a larger difference) [Kim, Dennick, 1970; Learmonth, 1975] (see Appendix B, Sequence 2a, b). The outcome of this testing is shown in Table 2. If, for example, one examines the result on the pre-instruction inventory test, the value of the test statistic is reported as .208. The probability that an observation this large could occur by chance alone, given that H_o is true, is .198. This means that H_o could be rejected at the 19.8% confidence level. This level of confidence is too low for most statisticians to feel comfortable about rejection (normally one does not reject below the 90% level of confidence). Note that only with the Prior Exposure Variable does the observed difference between group means approach statistical significance (i.e., 88.7%) and the data shows that this difference is in favor of the CLR group. Thus, it is felt that the assumption that the groups are matched is a valid one.

However, since it is known by statisticians that non-parametric tests lack power, "confidence" in statistical findings occurs if one is led to reject H_o, not accept it. While the desire here is to establish H_o, there are no known tests to do so. Nonetheless, armed with the non-parametric results it is worthwhile to subject the same data to the relevant parametric tests. If these tests also deny us, with sufficient

Table 2

Variable	Test Statistic	$P(D_{mn} = TS) = (1-P)$
Pre-Instruction Inventory Exam	.208	.198
Mental Maturity Language	.292	.558
Mental Maturity Non-Language	.313	.636
Personality Traits Personal	.229	.278
Personality Traits Social	.229	.278
Prior Exposure Mathematics Courses	.417	.887

These results indicate that H_0 could not be rejected across any variable at normal levels of significance employed by statistical researchers. Only across the "prior exposure" variable does the difference between groups approach statistical significance ($p = .113$) and that difference is in favor of the CLR group (see Appendix B, Sequence 1; 2a, b).

confidence, the ability to reject H_o, then even further evidence is presented to support the assumption of matched groups.

The next test employed was the parametric "student's $t_{(26)}$", or the univariate F-ratio $(_{1,26})$ (one-way Analysis of Variance, ANOV). This test also examines the difference in the observed means of groups across variables (one at a time). It requires two assumptions, in addition to that of independence between observations as required by the non-parametric, Kolomogorov-Smirnov (K-S) test. These are: a) that the underlying distribution of the random variable is normal and b) that the variance of the random variable is the same for both groups. These are strong assumptions, justified by common usage and the design of several of the measures of effectiveness. When they are satisfied the test has great power. Nonetheless, the results of this analysis are in agreement with the K-S test (see Table 3). This table provides two measures of the significance of the observed differences in group means across the variables (take one at a time). The first is Wilk's Lamda (λ), which varies between zero and one. As (λ) approaches one, the probability that the group means are disparate goes to zero. The other is the value of the test statistic $F_{(1,26)}$. For this statistic to have any reasonable significance it must be greater than one. Again, only across prior exposure is this the case. The observed value 2.5355 would enable one to reject H_o at some confidence level below 90% which is considered too low a level for rejection.

Table 3

Variable	Wilk's Lamda	$F_{(1,26)} = t^2_{(26)}$	Significance
Inventory	.9668	.8929	*
Mental Maturity Language	.9986	.0353	*
Mental Maturity Non Language	.9710	.7760	*
Personality Personal	.9936	.1669	*
Personality Social	.9925	.1957	*
Prior Exposure	.9711	2.5355	p .120

These results indicate that the difference between group performance across the variables is not significant except over the Prior Exposure where H_o could be rejected at p .120. The difference between group performance on this variable is in favor of the CLR Group.

(* implies that statistical significance at .10 level is not reached). (See Appendix B, Sequence 3).

Each of these tests examines the input variables singly. In order to look at the combined effects of all variables taken together, a two-way, nested, factorial ANOV was performed (Table 4). Again, the results support the hypothesis of equality of group means. Table 4 reveals that the "between students" effects are not significant and neither are the effects between groups. As might be expected there is an effect across variables (due to different scales of measure) and there is a cross effect indicating that neither group performed consistently better or worse across all variables (see Appendix B, Sequence 4).

Finally, in an effort to determine the existence of a discriminating capability across the student input characteristic variables and to take advantage of the power associated with a procedure allowing the step-wise inclusion of the variables, a technique known as discriminant analysis was employed [Klecka; in Nie et al, 1975]. The use of discriminant analysis in this manner is innovative and valid. All of the data was input to the program. The computer calculated the prior probability of a student being in one group or the other (based upon group size (m, n)). It also read the scores of each student on each variable. Armed with these facts it computed a discriminant function (see Appendix B, paragraph A4 for a detailed explanation) and attempted to classify students into their proper group. Failure to do this well would indicate little differences between the groups but the ability to do so better than expected (if the only difference between groups was, in fact, their size) would

Table 4

Source	df	SS	MS	EMS	$F_{.90}(v_1, v_2)$
Between Students_R	27	2864.33	106.09	.97	N.S.
Groups	1	10.71	10.71	.10	N.S.
e_{ir}	26	2853.52	109.75		
Within Students	140	136398.45	974.27	35.17	Sig.
Characteristic Variables	5	132059.48	26411.90	953.50	Sig
$(C \times G)_{ij}$	5	738.33	147.67	5.33	Sig
e_{ijk}	130	3600.64	27.70		
TOTAL	167	139262.78			

Source	df	SS	MS	EMS	$F(v_1, v_2)$
Students	27	2864.33	106.09	.11	N.S.
Error	140	136398.45	974.27		
TOTAL	167	139262.78			

indicate that the variables were discriminating, as desired. The results were most rewarding. It was found that significant differences did not exist between the groups but that students could be classified with 78.5% accuracy.

This compares favorably with the 57% accuracy expected if the variables were not discriminating at all, and the program was classifying based upon prior knowledge of group size (i.e., 16/28 = .57). A slight variation between the groups that might have had an impact was the different number of average credit hours in their course loads. The CLR group averaged 17.54 hours and the PSI group took a mean weekly load of 19.39 hours. This chance occurrence was a result of academic scheduling from the administrative offices of NPS. As a result three PSI students opted to drop MA 2045 prior to the end of the first week of instruction. If the CLR had been burdened with the additional hours it would have been necessary to make experimental adjustments. Since the PSI group had the disadvantage it was decided to permit the groups to continue in the investigation, thus enriching the potential for gathering information on the effects of PSI under the adverse conditions of a relatively heavy academic schedule.

One other source of possible variation, the effect of the professor, was controlled by having the same man assume responsibility for the conduct of both classes. He would teach the CLR group in the traditional manner and he would manage the PSI group according to the principles

of the Keller Plan. The professor (Alan B. Shorb, Ph. D) was chosen for the critical role because he had taught MA 2045, via CLR, many times and was considered particularly capable (according to opinion expressed by his previous students and his faculty peers as well). In addition he had attended a three-day workshop in PSI conducted in the winter of 1975 at NPS. The workshop had been led by Professor Ben A. Green, Jr. of the Institute for PSI, Georgetown University. Further, Professor Shorb was familiar with the content of Professor Weir's written PSI course materials and had expressed no a priori preference for teaching via one mode or the other. He had not taught MA 2045 via PSI prior to the experiment and it was felt that this fact would enable one to determine whether Professor Weir's course was truly systematized (i. e., could a competent professor teach successfully via PSI with course materials he had not worked with before). The same text [Steinberg, 1974] was used for both groups.

The tutors (two) were selected from the student body. One was an electrical engineering major in his 3rd quarter of an eight quarter program. He had previously taken MA 2045 via PSI from Professor Weir. The other was the author of this thesis, an Operations Research/ Systems Analysis major in the 7th quarter of an eight quarter sequence. He had taken MA 2045 in his 2nd quarter via CLR. Both tutors were thoroughly briefed on their responsibilities by Professors Weir and Shorb.

D. THE CONDUCT OF THE CLR CLASS

The class met for lectures at a pre-designated time for fifty minutes, three days a week (Monday, Wednesday and Friday from 1110-1200 hours). The professor gave reading assignments from the text, homework problems, and offered the students an opportunity to work on two outside projects (for extra credit to be applied to the FE score). He announced the objectives of the course on the first day of class and explained to students how they would be graded. The students were also told that they were the control group in an experiment but that this condition would in no way affect their instruction or their grade. They were given their pre-instruction tests and great care was exercised to motivate them to do their best. They were assured, for example, that their scores would not be publicized but that they could be counseled on an individual basis, if they so desired, as to the significance of their performance on each test (92% of the students desired such counseling and showed great interest in their scores).

Each class hour progressed in the traditional manner. The instructor lectured each day, answered questions from the class, and administered two "mid-term" examinations, each of which consumed one class period and contributed approximately one-fourth to the course grade. The FE, plus any extra credit, contributed approximately one-half to the course grade. These proportions are not exact as the professor reserved the right to modify any grade as his subjective evaluation dictated (only the

course grade was so affected; the tests and extra credit problems were graded as objectively and fairly as possible, with consideration for the awarding of partial credit on questions that were not answered perfectly). All tests were closed book with no notes available to the students.

E. THE CONDUCT OF THE PSI CLASS

This class was conducted, as closely as possible, according to the principles of the Keller Plan. All materials (study guides, diagnostic tests, etc.) had been prepared in advance, and all logistical and administrative considerations had been accounted for prior to the first day of instruction. The students were advised that they were the experimental group and would receive their instruction via PSI. They were offered the opportunity to withdraw from the group without prejudice if they so desired. None withdrew initially. The PSI technique was explained in great detail, both verbally and in a written(handout) "Course Policy Statement," and all questions concerning the conduct of the course were answered. The students met their tutors and were told that they could employ either of them as they so desired. It should be noted that one tutor could have sufficed for a class of this size (m = 12). However, it was felt that two tutors would enable the students to pick the man that they could most easily relate to, would reduce the requirement for the instructor to "teach" the PSI students, and would reduce the amount of time that a tutor was unavailable for consultation because he was grading a diagnostic examination.

The PSI students were given their pre-instruction tests and were given the same careful instructions and motivational speech as were their CLR counterparts. Once the course began the students progressed through the material at their own pace. When they had completed the work required on a unit in the sequence, they sought out a tutor and requested the diagnostic examination for that unit. They could do this between 0900-1030 hours on Mondays, Wednesdays and Fridays. All work on these exams had to be completed by 1100 hours. During these hours the instructor and tutors were present in a designated classroom and were available for consultation or test grading. At other times during the week tutors could meet with students at mutually acceptable times for consultation but not to administer tests.

After a student completed a diagnostic test he brought it to the tutor for immediate grading. The test was gone over thoroughly in the presence of the student and if the student received a score of 90% or better he was permitted to proceed to the next unit. Otherwise he was advised that he had not "mastered" the material and would have to retake a similar exam. Exams could be retaken the same day but only after a delay of thirty (30) minutes. (Since exams were usually forty-five minutes long and required approximately ten minutes to grade; this reduced the likelihood that two exams could be taken in one day. This was done in order to motivate students to thoroughly prepare for each exam and to master them the first time.) It should be noted that even if a student

received a score of 90% or better, any errors he made were discussed at length to ensure that he completely understood the material. If he failed to demonstrate mastery the first time, the second exam given was different than the first but similar in design. All exams were closed book and no notes were permitted in the examination room.

The professor presented three optional lectures during the conduct of the course. The information contained in these lectures was designed to motivate the students and to enrich their appreciation for the applicability of the information they were learning to the solution of practical problems. Students could attend these lectures regardless of where they stood in the course but were advised that their understanding of the lecture material would be a function of their progress to date.

As in the CLR group, the PSI students were advised of the extra-credit assignments available to them and that completion of these problems would enhance their FE scores.

The grading of the PSI group was quite different from that of the CLR group. At NPS an average grade of B (3.00) is the minimum for graduation with a Master's degree. Thus, it behooves each student to set his standards at that level, at least. Accordingly, the PSI group was told that if they completed all of the required units, mastering (90% or better on the diagnostic tests) each in turn prior to the end of the quarter, and if they took the final exam (scoring at least 90%), they would receive a course grade of A (4.00). Otherwise their course grade would be

diminished according to their performance on the final exam (a course grade of B, for example, required a score in the 75-79% range). The final exam could be taken at any time, so that a student who completed the course early would not have to wait to take it during "finals week." To give credit where credit is due, students who completed all units (recall this implies mastery) would receive at least a B- (2.66) even if they did not take the final or performed very poorly on it. This scheme was adopted because it was believed that a student who went through the entire course (taking thirteen exams along the way, scoring at least 90% on each one) deserved at least a B- (2.66). Moreover, a grade of B- was a slight penalty (in view of the B (3.00) minimum graduation requirement) which would encourage each student to take the final exam and perform to the best of his ability. The fact that every student took the final exam and demonstrated a high level of performance on it is supportive of the reasonableness of this kind of scheme for NPS students.

It is acknowledged that the self-pacing aspect of PSI is mitigated by the requirement that students complete the course in not more than twelve weeks. However, in this experiment practicality prevailed. A "quarter" is only so long and MA 2045 is a required course in these students' curricula. The school administration could not allow a student to continue the course beyond the end of the term without significant revisions in its curricular programs.

F. THE OUTPUT MEASURES OF EFFECTIVENESS (MOE s): PSI vs CLR

1. Percent of students completing the course

As the course began in earnest, nine of the original twelve students in the PSI group remained. The three students who dropped the course, in its first week, did so for personal reasons relating to the large academic loads they were scheduled to take and their own evaluation of how they could best meet the requirements of NPS for graduation. One of the students had been away from higher education for twelve years, for example, and feared that devotion of the time required by a PSI course would inhibit his ability to get an acceptable grade average in his other courses. A second student had taken an advanced course similar to MA 2045 in undergraduate school (USNA) and felt he could drop MA 2045 in his first quarter with the least effect on his ability to satisfactorily progress through his other courses. The third student advised his professor that he wished to "audit" the course, that is, perform the required unit exercises at his leisure and avail himself of the tutors, but not receive course credit. This would obviate the necessity of taking the time consuming examinations and probably ensure him of a high grade when he did take the course for credit at a later time. Thus, no PSI student dropped the course due to dissatisfaction with the mode of instruction. All of the nine students who actually began the course were able to complete it well within the time constraints.

All sixteen CLR students completed the course for credit. Thus, across this criterion the two groups performed equally well.

2. Final Exam Scores (FE)

The FE was a two-hour closed book comprehensive test that covered all the material presented during the quarter. It was written by the course professor, in collaboration with the Chairman of the Mathematics Department, and reviewed by the tutors and other mathematics professors. There were two similar versions. The first was given one week early to those PSI students that completed the course prior to the end of the quarter. The second was given to the remaining PSI students and to all of the CLR students at the scheduled final exam period designated by the academic offices. Both tests were administered and graded by the tutors of the PSI course. The same grading scheme was used for both groups. Afterwards the tests were reviewed by the professor. He lowered no grades but reserved the right to give additional credit where he found the tutors had been too strict. Ironically, this occurred on two papers belonging to PSI students.

The performance of students on the FE are summarized in Appendix B, Sequence 5 a, b. An analysis of the results shows that the PSI students scored significantly higher than the CLR students ($p = .005$).*
Thus, it can be infered that the group means are statistically different and, since all other relevant variables impinging upon group performance

---

*Recall that the significance level is the probability of obtaining differences in the group means as large as those observed, or larger, due to chance alone, when the group means are assumed to be equal.

were held constant (except the form of instruction), that PSI produced the significant difference observed. The test used to obtain this significance level was the non-parametric Kolomogorov-Smirnov test described earlier [Kim, Jennick, 1970]. As a matter of interest the data was also analyzed according to the univariate $F_{(1, 23)}$ test. This test revealed the difference to be significant also ($p = .004$). Thus, if the more rigorous assumptions of normality and homoscedasticity are made, then the more powerful parametric test rejects the hypothesis that the mean scores are equal with a high level of confidence (i. e., 99.6%).

3. <u>Course Grades (CG)</u>

The CG for both groups of students is not independent of the final exam (FE) score. In the case of the CLR group two "mid-term" exams, of fifty minute duration contributed approximately 50% to the grade with the FE contributing the remaining 50%. However, the performance of the students in their classroom meetings weighted on the professor's subjective evaluation of each individual. The exact nature of the professor's subjective input to the CG of the CLR group cannot be measured. However, in no case did the professor's intuitive input to the course grade result in its being lowered below that analytic computation described below:

$$CG_{CLR,\,i} = .25\,(MT_1) + .25\,(MT_2) + .50\,(FE)$$

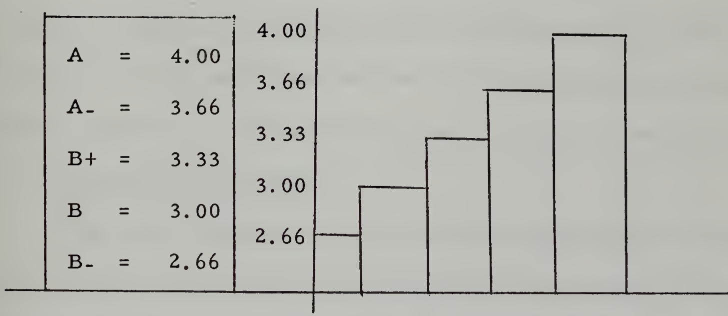
where

CG_{CLR} = course grade of CLR student (i); $i = 1, 2, \ldots, 16$

MT_j = mid-term exam grade on exam $(j)_j$ $j = 1, 2$

FE = final exam score.

In the case of the PSI Group the CG was totally objective. If a PSI student completed all thirteen required units he merited, automatically, a CG of B- (2.66). If he chose to take the final exam (FE) he could improve his CG according to the scale described below:



Below 75% means a B-   70   75   80   85   90   100   % correct on FE.

The performance of the students as measured by CG is presented in Appendix B, Sequence 5 a, b. An analysis of these results according to the K-S test shows that the mean grade of PSI students is numerically, but not significantly, higher than that of the CLR group ($p = .146$); but according to the univariate F-ratio test this difference is statistically significant at $p = .025$.* The fact that the CLR group mean

*It is a rather common occurrence for non-parametric tests to judge differences in observed means to be less significant than parametric tests judge the same differences. This is so because parametric tests have greater power derived from their reliance upon assumptions of a particular distribution to the data and homoscedasticity. If one is willing to make the assumptions necessary to employ the parametric test then H_o can be rejected at the .975 confidence level.

score on CG is closer to the PSI group than was their mean FE score is explained by the boost that some CLR students received as a result of their extra credit work and the subjective evaluation of the course professor. It is not felt that this subjective evaluation, where applied, was unjust. However, it is apparent that the PSI group, as a whole, would have appeared even more impressive vice the CLR group had no subjective input been added to the formulation of the CLR students' CG.

4. Rate of Progress (RP)

The rate of progress is a variable that represents the average number of units that a PSI student completed per "work day" (six work days per week) throughout the course. Its tabulation here is intended to illustrate the effect of the self-pacing aspect of PSI. For comparison purposes it is assumed that all CLR students progressed at the same theoretical rate during the quarter. Obviously this assumption ignores the fact that not all CLR students "learned" the material equally well, that some were pushed too fast while others were, perhaps, "held back." It further ignores the fact that the professor did not spend equal amounts of class time on each topic within his course. However, it was not practical or even possible to get a measure of the actual RP of the CLR group and the assumption of linear RP must suffice for comparative purposes.

The results show that all of the PSI students completed the material not later than the eleventh week of the quarter. Five of them completed the course (excluding the FE) in the tenth week. A graph

depicting the relative rates of the two groups is included in Appendix B, Sequence 6. It is of special interest that the CLR group (when interviewed) indicated that their actual rate of progress was not linear at all. In fact they spent too much time on mechanics early in the course and were forced to rush through more difficult theoretical material presented towards the end of the quarter. This revelation, while perhaps disturbing, is evidently quite common in CLR classes. The phenomenon is attributable to perhaps several variables in a CLR course, one of which is its absolute reliance on the professor as a primary source of information, as well as an interpreter of the textual materials. Therefore, in a CLR course, the professor sets an "average pace" that seems to him realistic for "most" of the students enrolled in the course: a pace that allows for enough time to cover all of the critical material as required by the course syllabus and yet avoids spending too much time on material that is not particularly difficult. However, this is never a trivial task, and if the professor is concerned about student abilities to satisfactorily understand the more difficult course materials he may feel compelled to spend more time on the earlier fundamentals. On the other hand, in a PSI course, the stress is on the <u>written</u> word so the student can, if he wishes, plan his progress from the first day. If the student needs to spend considerable time on the fundamental ideas at the beginning (because of deficiencies in his background work) he has the opportunity to do so (with the aid of the textbook, study guide, and so forth which

are all carefully prepared in advance, as well as any individual aid he may require of the professor), and this is not at the expense of other students taking the course for whom many of the course fundamentals are of a review nature (thereby leaving more time for these students to study the more difficult materials and, perhaps, finish the course in a lesser amount of time).

The variable RP of each student in the PSI class can present logistical and administrative problems. For example, all course materials must be available when a student is ready for them. It is, therefore, incumbent upon the course professor to ensure that the course is intact at the beginning or that he is well ahead of his most advanced students. Further, it is possible that very good, highly motivated students with few outside distractions may complete a PSI course, even a difficult one, very quickly. This has great advantages if the student is enrolled in a flexible curriculum. However, in most institutions, operating on the concept of an academic year, new courses are not available to students at arbitrary times during the term. It is well for the administration to consider making allowances to accommodate exceptionally gifted or unusually slow students. In order to assist in planning and predicting student completion times the reader is referred to the work of Wagner, et al [1973]. In that paper the authors develop predictive models that provide indicators of the probable length of time a given student will require to complete an individualized course of instruction. These

indicants can be used to help program students for future courses independent of calendar academic years.

5. Facilitation

This output measure attempts to reflect the effects, if any, of PSI on the performance of students in their other CLR classes. It is of importance to know whether PSI, with its high demand on student body time, forces its pupils to pay a price (diminished performance in their other courses) for their accomplishments in the PSI course. It was felt that this effect might be magnified within this experiment because the PSI group was taking more credit hours, on the average, than was the CLR group (i.e., 19.39 hrs vs 17.54 hrs). In addition several of the PSI students were taking another course via PSI, FORTRAN PROGRAMMING. The typical PSI student's course load vice that of a typical CLR student is presented in Appendix B, Sequence 7a. As a result of the heavy demands upon their time one might expect to see slightly lower cumulative grade point averages (GPA) for the quarter, within the PSI group than in the CLR group. Therefore, it is surprising and significant that the PSI students outperformed their CLR counterparts, not only in MA 2045, but in their other courses as well.

The grade point average (GPA), with the effect of MA 2045 removed, of the PSI group exceeds that of the CLR group by a margin that is significant at ($p = .074$) according to the K-S test and at ($p = .075$) according to the univariate $F_{(1, 23)}$ ratio, see Appendix B, Sequence 7b.

This finding is extraordinary because the PSI and CLR groups were statistically identical over the relevant input variables of Mental Maturity, Personality (Motivation), Prior Exposure, etc., which were measured before any instruction at NPS began. These findings are in conformity with those of Schimpfhauser, et al [1974] at Ohio State University College of Medicine. In addition it suggests that PSI, with its requirement for mastery and its logical step-by-step progression, may not only imbue in the student a sense of achievement and a desire for high standards but, furthermore, it may teach him techniques of study that are more efficient and more effective than those learned via CLR. Moreover, the PSI student may develop a higher degree of testmanship than he held previously. After all he must take at least one test per unit, he must learn to work quickly and with great accuracy, and he must certainly dissolve any damaging fears of tests that he may have harbored previously.

6. Attitudes

Finally, the output measure that attempts to reveal something of the opinions of the students themselves is considered. Several previous studies conducted in this regard report that PSI contributes to "positive" attitude growth [Kulik, 1975; Stice, 1971; Gallop, 1970; Green, 1971]. What exactly is meant by positive attitudes varies among studies but one of the clearest representations of this concept is explained by Silberman and Parker [1974]. In their report they concentrate entirely

upon the measure of attitudes among students after exposure to a course in organic chemistry at the undergraduate level. The basic tool they employed consisted of a questionnaire that was carefully designed and checked for internal consistency, reliability, and validity. Their results indicate that overall the attitude of their PSI students toward their course and their mode of instruction was significantly more "positive" ($p = .05$) than the CLR group. This result was observed in spite of the fact that both groups were statistically equivalent across seven student characteristic variables at the beginning of the course and that the same professor taught/managed both classes.

In order to measure student attitudes in this experiment a questionnaire similar in scope and construction to that of Leiberman's and Parker's was designed to be administered to both PSI and CLR students in MA 2045. The questionnaire (see Appendix B, Sequence 9b) was carefully explained to both groups and was given to them ten days before the final exam. They were asked to respond to the questionnaire at their leisure, to consider each question carefully, and to return the questionnaire not later than the day of their final exam. Twenty-two of the twenty-five students (88%) responded. Response was voluntary and no names were attached to any of the questionnaires, only the group symbol was affixed to ensure correct classification of them for purposes of analysis. The results of this analysis are presented in Appendix B, Sequence 9a. Note that a score close to zero (0) reflects a more positive

attitude, the maximum negative attitude score being five (5). It is apparent that across all forty-three responses to the questionnaire the two groups are nearly identical. The PSI group exhibits a more positive attitude on 74% of the questions but the difference reaches statistical significance (p = .10) only 21% of the time (9 of 43). However, in every significant case except one (question number 23) this difference is in favor of the PSI group.

When all the scores of the questionnaires were subjected to discriminant analysis it was learned that the matrix of correlation coefficients was too ill-conditioned to allow for the construction of an inverse (the determinant of the matrix was near zero). This was because many of the students possessed attitude vectors which were nearly parallel. To overcome this difficulty and get to the root of the attitude question, the mean scores of the groups across the nine questions (where the attitudes were determined to be significantly different: those questions marked ** in Sequence 9, a, b of Appendix B) were analyzed by program DISCRIMINANT. The outcome resulted in totally accurate classification of students into their respective groups. That is, based upon a student's response on the nine questions alone, the computer was able to predict, with 100% accuracy, the actual group to which the student belonged.

The conclusion is that the PSI mode of instruction generated within its students certain attitudes that tended to be more positive than those exhibited within the CLR group.

A general summary of the attitude questionnaire results indicate that students of MA 2045 under the PSI mode were required to assume more active responsibility for their own learning, learned more from the course, could better cope with the amount of information presented, enjoyed sharing their knowledge with their classmates to a greater degree, looked forward to applying their skills to the solution of practical problems to a greater extent, and perceived more clearly the objectives and goals of the course.

Of additional interest might be the feelings and attitudes of the professor and tutors of the PSI course. These men felt that the course had been a total success, not only across the objective MOE's but from the point of view of the personal satisfaction they derived from their association with their students. The professor indicated he enjoyed his role in both courses. He lectured more via CLR but his role was more tutorial and highly individualized under PSI.

7. Summary

If one considers together only the output measures which are devoid of subjective components, that is, FE, CG, and GPA (with the effect of MA 2045 removed) the picture presented is a theoretically balanced meshing of student performance measures that have been part of traditional education for the last half-century. These three measures were analyzed singly in earlier sections. It was also considered appropriate to examine their combined effects. First, through Analysis

of Variance (ANOV) and then via discriminant analysis. The results of the ANOV are included in Appendix B, Sequence 8. They reveal a significant difference across the variables at the $p = .01$ level, "between the groups." This indicates that the hypothesis that the two groups performed equally well can be rejected with a high level of confidence: The PSI group did display greater ability and experienced a marked facilitation effect as a result of their exposure to PSI. The observed significance of the cross effect (that is, output measures across groups) resulted from the varying degrees of performance difference demonstrated within each group across the variables.

When the output measures (FE, CG, and GPA) are subjected to discriminant analysis the computer is capable of properly classifying students into their respective groups 80% of the time. If the groups were essentially equal, we would expect the computer to classify properly only 64% of the time (this is because the prior probability of a student being from CLR is $.64 = {}^{16}/_{25}$, based upon the size of the CLR group).

It is clear that the PSI group demonstrated superior performance across all output measures (MOE's). These findings are in complete agreement with those of the previous studies mentioned earlier and lead to the conclusion that PSI, as conducted at NPS during this experiment, is an effective method for teaching MA 2045 and is probably more beneficial for students than is the CLR technique currently employed.

IV. DISCUSSION - EXTRAPOLATION - FURTHER STUDY

It is apparent that the experiment, as designed and conducted, was successful: it produced meaningful results and accomplished its purpose. The experiment leads one to conclude that PSI is an effective and efficient method for teaching MA 2045 at NPS. It is efficient in that its students progress, on the average, more rapidly and with better results than the standard CLR students. It is effective as evidenced by the facilitation effect produced and the demonstrated performance on the final examination. To discover these things is important in itself. However, at least three questions remain unanswered with respect to PSI at NPS:

1. Will students who take PSI courses be able to transfer their knowledge to other related courses with relative ease?

2. Will students who learn via PSI retain the information they assimilate for a longer time period than those who learn via CLR?

3. What other courses, besides MA 2045, are adaptable to PSI and on what scale can PSI be taught at NPS?

Questions 1 and 2 could be answered by conducting another experiment: One could simply study the transfer and retention effects of PSI on the same students used in this experiment. Several studies [e.g., Corey and McMichael, 1971; Moore, J. W., et al, 1973] already attest to the noteworthy benefits students derive in these important areas and there

is little reason to suspect the same results would not accrue to NPS officers.

The third question, however, deserves some special attention. Imbedded within it is an implicit issue: Should or could an entire curriculum (or some percentage thereof) be taught via the PSI mode? Only time and experience can accurately provide the answers. There are, however, some important indicators of probable results and potential benefits to be derived from an attempt to learn the truth of the matter concerning PSI. The interested decision maker would be well advised to read the admonitions contained in Green's "Fifteen Reasons Not to Use the Keller plan," Sherman's "PSI Some Notable Failures," and Greenspoon's "Should An Entire College Curriculum be Taught by the Keller Method" [All in Sherman, 1974]. Contained in these papers are the hard facts gained from many trials by dedicated men of various disciplines in their efforts to hone PSI to its present edge. Even they admit that the full potential of the Keller system is yet to be realized. Of special note is the fate of PSI at the hands of Massachusetts Institute of Technology (MIT) [Friedman, et al, 1975]. This paper describes the decision to discontinue PSI in Physics. In spite of small scale successes in introductory physics courses, the transition to large scale teaching was mismanaged. The endeavor was discontinued after the fourth trial and large-scale physics instruction reverted to the traditional CLR mode. It is not the purpose here to summarize the Friedman report. What is

necessary is to point out the importance of detailed management planning prior to the adoption of large scale PSI. In fact, the importance of such planning cannot be overemphasized.

One must examine the nature of the course(s) to be taught. The literature is replete with successes in many disciplines (e.g. Mathematics, Chemistry, Physics, Psychology, Electricity and Magnetism, etc.). However, the courses have largely been "introductory" in nature. This means that they have dealt with the basic principles, mechanics, language, and foundations of the discipline. There is no reason to believe that "advanced" courses cannot be taught via PSI, but since no PSI course can succeed without the presence of qualified and dedicated tutors, if in an advanced course the tutors are incapable of doing their job because their expertise is not up to the task, then the students cannot turn to them for the meaningful counsel they will require.

Given that the subject matter is teachable via PSI, and that a professor and tutors are available, then the next step is to write the PSI course itself. This procedure is not trivial, requires the existence of a gifted and very knowledgeable author, and can be time consuming. This man must be thoroughly familiar with PSI, its principles, objectives, and characteristics. Once the course is written it must be validated, at least on a small scale. (The course that is the subject of this paper was presented twice prior to the experiment and was revised to improve it after each trial. The course matured with age and was taught by three

different professors and six tutors during this maturation/validation process. It is without question now ready for presentation to large groups.) Additional work must still be accomplished prior to implementation. Materials must be printed, tutors must be obtained and trained, and a professor who is totally familiar with his role as manager of a PSI course, must be selected. If this sort of preparation is made with additional management planning, then large-scale teaching of PSI can follow. It will provide its students with all of the benefits experienced by those of smaller groups and the same sort of efficiency and rate of progress curves would be expected.

If, in fact, an entire curriculum is to be taught via PSI then a complete rethinking of the nature of the "academic year" is in order. Since students will finish their courses at varying times, they should be able to continue with the next course in sequence without having to wait for the end of an arbitrary "quarter" to end. Careful consideration would also need to be given to the number of courses a student should take simultaneously: PSI courses require intense and diligent study. If too much is expected of a student he may tend to do less than if he perceived his work load to be manageable and realistic. This thesis does not address these issues. However, it is obvious that such matters must be carefully considered prior to any decision to expand or extrapolate PSI at NPS beyond the "small-scale" level.

V. SUMMARY AND CONCLUSIONS

A carefully controlled experiment was designed and conducted for the purpose of determining the relative instructional effectiveness of PSI at NPS. It was felt that the results of the experiment would provide valuable information pertinent to the educational processes managed under DOD authority in general. The vehicle for the study was MA 2045 (Computational Linear Algebra). The course was written in its PSI form by Professor Maurice D. Weir. Prior to the conduct of this experiment the course was tested twice and carefully revised following each trial.

Once it was subjectively established that the PSI course was valid, two groups of students ($m = 12$; $n = 16$) were selected at random and taught MA 2045 via PSI and CLR. The same professor conducted both classes. Both groups were screened, via testing, prior to any instruction at NPS to lend credence to the hypothesis that both groups constituted independent, random samples drawn from the same population.

The two groups were then given the linear algebra course. The output variables, (MOE's) were:

1. percentage of students completing the course
2. final examination scores
3. course grades
4. facilitation
5. rate of progress
6. student attitudes.

These output variables were subjected to several statistical tests (parametric and non-parametric) and a sophisticated data analysis technique (Discriminant Analysis) in order to determine the significance of the distance of the group centroids in reduced space.

Group performances across variable one (above) were equivalent. Across the other variables, however, PSI students demonstrated a statistically significant superiority compared with CLR students in MA 2045 at NPS under the conditions described in this paper. As a result PSI is considered more effective and more efficient than CLR for teaching the course Computational Linear Algebra, MA 2045. It is considered valid to extrapolate beyond this specific course and to apply the results of this study to similar courses of similar scope in other areas of mathematics and other disciplines as well (see Appendix A). If, however, the results of this study are to be extended to a larger scale, then it is strongly recommended that the admonitions and considerations explained in the discussions section of this report be taken to heart. If an entire curriculum is to be constructed around PSI then careful analysis of the appropriateness of the structure of the traditional academic year is warranted.

It is recommended that further study to determine PSI's effect on retention and transfer be conducted. Finally, the relative cost-effectiveness of PSI for NPS should be documented.

APPENDIX A

Student Performance in a PSI Course in Networks Flows and Graphs (OA 4633)

It is the position of adherents of PSI that the teaching mode is valid for many types of courses, at all levels across a broad spectrum of academic disciplines. The main body of this report discusses an experiment designed to measure PSI's effectiveness vice CLR and employed MA 2045 as a vehicle. Concurrently a course in Networks Flows and Graphs (OA 4633) was being taught by Commander Joseph Cyr, Curricular Officer, Operations Research, Systems Analysis. Commander Cyr, like Professor Shorb, had attended a workshop on PSI and had been interested in converting his Networks course for presentation via PSI. By Quarter IV FY 75 his materials were ready and he permitted his students (12) to choose the mode of instruction (PSI or CLR) they preferred. They were all MS (OR/SA) candidates in their seventh quarter of an eight quarter sequence. They had signed up for OA 4633 as an elective.

Eight of the students opted for PSI and four for CLR. Prior to instruction the two groups were compared across their Grade Point Averages (GPA), Mental Maturity scores and Personal Inventory scores. Although with small sample sizes, statistical inference has little meaning, the group means were very close across all variables (K-S, non-parametric test of equality of group means and suggested that there was no reason to reject the hypothesis of co-location of group centroids.

Commander Cyr selected tutors (2) for his PSI group and conducted his course with these students according to the principles of the Keller Plan.

At the completion of the course the students were compared across the output measures generated by the final exam, the rate of progress, and the student attitude questionnaire. The results are in complete agreement with those found in the main experiment discussed in the body of this report. Again, it is not meaningful to infer statistical significance with such small samples. However, the success of the course and the impressive performance of the PSI students confirms the viability of teaching OA 4633 via the Keller Plan.

This fact supports the position taken in the body of this paper that PSI will work effectively at NPS on courses other than MA 2045.

APPENDIX B

Synopsis of Data Analysis

This section of the thesis is intended to epitomize the data collection, analysis and results of the experiment conducted to determine the relative effectiveness of PSI vice CLR for teaching MA 2045 at NPS.

A. PRIOR TO THE INSTRUCTION

1. After two groups of students were scheduled for instruction via one mode or the other (selection determined by NPS administrative offices against internal scheduling criteria) it was necessary to conduct certain tests to support the assumption of the existence of matched groups. A total of twenty-eight subjects were available for the purposes of this experiment. Twelve students constituted the PSI group ($m = 12$). The remaining sixteen individuals comprised the CLR group ($n = 16$). Measures of six student characteristics were obtained according to the methodology discussed in the main body of this thesis. The raw data is presented in Sequence 1.

2a. The next step was to investigate, using statistical inference, the hypothesis that the two groups constituted random independent samples drawn from the same population. Consider the possibility of characterizing each student according to the six dimensional vector constructed from his performance on the characteristic tests described in the main body. In general an individual student is described by the vector:

$$X_i = \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_j \\ \cdot \\ \cdot \\ \cdot \\ x_6 \end{bmatrix} ; \quad i = 1, \ldots, 12, 13, \ldots\ldots, 28$$

The set $\{x_i; i=1, \ldots., 12\}$ constitutes a data base as does the set $\{x_i; i=13, \ldots, 28\}$.

The space is of dimension six and the hypothesis to be tested states that any observed differences in the location of the centroids of the two groups is so small as to have reasonably occurred by chance alone. The first "test" of the hypothesis (H_o) consisted of the subjective inspection of the historgrams generated by plotting the frequency distribution of each component of the vectors (x_i) singly. Since these histograms all assumed the same shape, over their respective ranges, it seemed reasonable to continue with another test.

The test selected was the K-S non-parametric two-sample test of the homogenity of cumulative distribution functions (one-way test). This test was applied across each component of the vector sets one at a time. The results of this analysis are portrayed in Sequence 2a.

2b. Since the power of any test is realized when rejection of H_o is possible, at an acceptable level of confidence, it was decided to test the null hypothesis with a more powerful parametric test. This test, the univariate $F_{(1, 26)}$ ratio required the assumptions of a normal distribution underlying the random variables (X_{ij}) and equality of variance within the samples across the variables. If either of these assumptions are invalid then the reliability of the test is questionable. However, it is known that the F-ratio is robust with respect to these conditions and since human performance is often normally distributed the test was employed. Sequence 3 displays the outcome of this analysis.

3. Both the non-parametric analysis of Sequence 2a and the parametric investigation of Sequence 3 considered the six components of the student characteristic vector one variable at a time. In order to look at all variables taken together a 2-way, nested factorial ANOV was performed on the data. The results are presented in Sequence 4.

The assumptions which are made when employing ANOV are the same as those of the univariate F-test (i. e., independence, normality, homoscedasticity). Fortunately, ANOV is also robust with respect to these assumptions and its results are in complete conformity with that which was expected, based upon the results of the previous testing. We find that there is no significant effect between students within groups and there is no difference between the groups themselves. The significant difference between variables (characteristics) is due to the various scales of the scores and

the cross effect is due to the fact that one group did not perform consistently better or worse than the other across all measures of effectiveness.

4. The final investigation of the data involved the use of a technique called linear Discriminant Analysis (DA). This technique is employable when one assumes that the discriminating variables have a multivariate normal distribution and that they have equal variance-covariance matrices within each group. In practice, the technique is very robust and these assumptions need not be strongly adhered to. Discriminant analysis begins with the desire to statistically distinguish between two or more groups of cases. In order to accomplish this task discriminating variables that measure characteristics of the groups are selected. The mathematics of the method are such that the variables are weighted and combined (linearly) so that the groups are <u>forced</u> to be as distinct as possible. The "discriminant" functions are of the form:

$$D_i = d_{i_1} Z_1 + d_{i_2} Z_2 + \dots + d_{ip} Z_p$$

where:

D_i : the score on discriminant function i

d_{i_j} : weighted coefficients

Z_j : standardized values of the variables.

In this analysis one such function could be constructed since only two groups were examined.

Once the function has been constructed <u>analysis</u> and <u>classification</u> can proceed.

In the analysis several tools are available to interpret the data; that is, to test for the success with which the variables have discriminated. The function itself is an axis in the variable space (six-dimensional) and it is used to study the spatial relationship between the groups. The coefficients are analogous to those derived in multiple regression or factor analysis. They serve to identify those variables which contribute the most to differentiation along the discriminant axes.

In the classification phase one attempts to place students into their respective group based upon their function score. If the variables are valid discriminators then we would expect proper classification a "high" percentage of the time.

The results of the discriminant analysis of the two experimental groups across the six variables of classification reveals successful differentiation is achieved; in that, proper classification of students into groups takes place with 78.57% accuracy. The overall significance of the separation of the group centroids allows for rejection of H_o with 95.3% confidence. This level is achieved principally because of the variable: "prior exposure." This score was derived from the algorithm below:

$$PE_j = \sum_{i=1}^{m} \left[(C_i - w(N_i) \right] + S$$

where

PE_j : the prior exposure score for the j^{th} student

C_i : a number assigned to a particular course which the j^{th} student took prior to matriculation at NPS

W : a weighting factor

N_i : the number of years since the j^{th} student took course i

S : a scaling factor to ensure PE_j is greater than or equal to zero

For example, suppose the j^{th} student had taken college algebra ten years ago, linear/matrix five years ago and linear programming four years ago. His PE score is derived as follows:

college algebra is worth one point
linear/matrix algebra is worth three points
linear programming is worth six points
W is equal to one-half; S equals five

thus

$$PE_j = [1-1/2(10)] + [3-1/2(5)] + [6-1/2(4)] + 5$$

$$= 5.5$$

It is felt that this algorithm gives an indication of the relative PE of the students but it is arbitrary and its validity is questionable. To measure the effect of this variable (taken singly its $F_{(1, 26)}$ ratio is 2.535, significant at p=.120) on the discriminant analysis as a whole, it was removed from the analysis. When this is done the significance of the distance between the group centroids drops markedly. The probability of error if H_o is rejected goes from (p = .047) to (p = .671). Thus, if PE is removed from the analysis the two groups are practically indistinguishable. The percentage of students properly classified drops to 64%, which is close to that which should be expected due to the computer's knowledge (a priori) of the sizes of the two groups (i.e., 16/28 = .57).

In any event the average PE score for the CLR group exceeds that of the PSI group. As such, if any advantage is presented it exists within the control group and not with the experimental group. This fact, coupled with the knowledge that the PSI group labored under a larger credit load than the CLR group forces one to view the results of the experiment as indicating marked advantages for PSI as a teaching technique over CLR as measured along the output variables.

B. POST-INSTRUCTION ANALYSIS

1. Final Exam Score: the performance of the nine students in the PSI course and the sixteen students in the CLR course is presented in Sequence 5a. Sequence 5b reveals the significance of the difference in group means according to the K-S test and the univariate F-test.

2. Course Grade: the performance of the students by group as measured by course grade is presented in Sequence 6a. Sequence 6b reveals the significance of the difference in group means according to the K-S test and the univariate F-test.

3. Rate of Progress: the average rates at which the two groups progressed through the course material is depicted in Sequence 6. The dots which have been plotted display the graph of the average number of units passed by the PSI group against days available in the quarter. The dashed line is the theoretical rate at which the CLR students, as a group, progressed through their course. The solid curved line shows the rate at which the CLR students perceived their group rate of progress. Note

how shallow the slope of this curve begins. Towards the end of the quarter it becomes steep, indicating the presence of a rush to complete syllabus material prior to the final exam.

4. Grade Point Average (with contribution of MA 2045 removed): The typical course loads, with associated credit hours, taken by the two groups is shown in Sequence 7a. The grade point average (GPA) achieved by students within the two groups were analyzed according to the K-S test and the univariate F-ratio. The results are summarized in Sequence 7b.

5. ANOV: The model used to examine the output measures FE, CG and GPA taken together is presented in Sequence 8.

6. Attitudes: Sequence 9a displays the different average attitude scores computed for each group on each question of the attitude survey. The range of scores is [0, 5] with a lower score indicating a more positive attitude. Sequence 9b presents the exact average score of each group on each question.

If one looks at the results of this survey across all questions taken together the conclusion that the two groups are indistinguishable is supportable by discriminant analysis. However, when one restricts attention to those nine questions (**) where the difference is so large as to be significant at the 90% confidence level it is found that the two groups are completely separable via discriminant analysis. That is, the students can be properly classified into their proper group with

100% accuracy. Of these nine questions only one (question #23) reveals a more positive attitude accruing to the CLR group. The conclusion is that the attitude of PSI students is at least as positive as CLR students but that across eight specific points it is demonstrably superior.

Sequence 1

Data Obtained from Tests to Measure Characteristics of the Student Groups

		VAR001	VAR002	VAR003	VAR004	VAR005	VAR006
PSI Group	S_1	0.0	47.0	47.0	85.0	83.0	3.0
	S_2	1.0	48.0	50.0	80.0	72.0	3.0
	S_3	2.0	50.0	37.0	52.0	47.0	3.2
	S_4	2.0	43.0	55.0	77.0	76.0	4.0
	S_5	2.0	49.0	48.0	63.0	72.0	4.6
	S_6	2.0	55.0	49.0	59.0	58.0	4.4
	S_7	3.0	53.0	43.0	68.0	55.0	3.8
	S_8	4.0	48.0	49.0	82.0	65.0	3.8
	S_9	5.0	53.0	43.0	66.0	60.0	4.7
	S_{10}	8.0	53.0	54.0	85.0	81.0	5.2
	S_{11}	12.0	53.0	49.0	81.0	79.0	5.3
	S_{12}	13.0	51.0	51.0	78.0	77.0	7.7
	$\overline{X}_i/S_i$	4.50 / 4.27	50.25 / 3.41	47.91 / 4.98	73.00 / 11.01	68.75 / 11.55	4.39 / 1.30
CLR Group	S_{13}	0.0	47.0	45.0	47.0	67.0	2.2
	S_{14}	1.0	50.0	48.0	77.0	69.0	3.8
	S_{15}	1.0	44.0	40.0	77.0	71.0	3.6
	S_{16}	1.0	54.0	40.0	82.0	64.0	4.6
	S_{17}	1.0	43.0	44.0	78.0	69.0	4.2
	S_{18}	1.0	51.0	43.0	78.0	75.0	4.8

Sequence 1 (continued)

		VAR001	VAR002	VAR003	VAR004	VAR005	VAR006
CLR Group	S_{19}	2.0	41.0	54.0	58.0	53.0	4.2
	S_{20}	2.0	50.0	36.0	69.0	56.0	5.4
	S_{21}	3.0	51.0	48.0	66.0	66.0	4.2
	S_{22}	3.0	49.0	50.0	82.0	78.0	6.4
	S_{23}	3.0	57.0	46.0	68.0	72.0	5.6
	S_{24}	5.0	53.0	53.0	89.0	77.0	6.0
	S_{25}	6.0	50.0	37.0	82.0	73.0	8.2
	S_{26}	7.0	55.0	46.0	82.0	77.0	7.8
	S_{27}	7.0	55.0	53.0	82.0	77.0	6.6
	S_{28}	9.0	59.0	54.0	78.0	82.0	7.6
	$\overline{X}_i$	3.25	50.56	46.06	74.68	70.37	5.32
	$/S_j$	2.72	5.01	5.87	10.68	7.91	1.68

Sequence 1 (continued)

VAR001: The score obtained by a student on the linear/matrix algebra pre-instructional inventory exam.

The range of scores = [0, 20].

VAR002: The mental maturity language proficiency score

Range = [0, 60]

VAR003: The mental maturity non-language proficiency score.

Range = [0, 60]

VAR004: Personality measure (personal traits)

Range = [0, 80]

VAR005: Personality measure (social traits)

Range = [0, 80]

VAR006: Prior exposure score

Range = [0, 10]

Sequence 2a

Variable	Test Statistic	$P(D_{mn} = TS) = (1-P)$
Pre-Instruction Inventory Exam	.208	.198
Mental Maturity Language	.292	.558
Mental Maturity Non-Language	.313	.636
Personality Traits Personal	.229	.278
Personality Traits Social	.229	.278
Prior Exposure Mathematics Courses	.417	.887

This table illustrates the great similarity in the performance of the two groups across the input characteristic variables. If one looks at the mental maturity language score, it is seen that the test statistic is .292. The probability that the K-S criterion variable could be greater than or equal to this value is .558. This means that if we were to reject the hypothesis (H_o) of equality of the two groups across this variable (VAR002) in favor of an alternative (H_1:PSI≠CLR) the probability that we would err in doing so is (1-.558)=.442. Thus, we cannot reject H_o with any confidence greater than 55.8%. This is far too low for most statistical/experimental work. Sequence 2b is a computer program which returns the test statistic reported and the exact probability of observing a value

as large or larger than that which was observed. In order to utilize this program the user provides the data and appropriate commands to read the input and write the output in the desired format. The program was written from instructions in Kim and Jennrich [1970] and adapted for use in this thesis by the author and Gerald Learmonth of NPS.

```
      FUNCTION CMN(X,M,Y,N)
      DIMENSION X(1),Y(1),Z(200,2)
      MN=M+N
      MNP=M*N
      CALL SORT(X,M)
      CALL SORT(Y,N)
      DO 1 I=1,M
      Z(I,1)=X(I)
    1 CONTINUE
      DO 2 I=1,N
      K=I+M
      Z(K,1)=Y(I)
    2 CONTINUE
      CALL SORT(Z(1,1),MN)
      XA=0.0
      YA=0.0
      K=1
      J=1
      DO 5 I=1,MN
      IF(Z(I,1)-X(K))6,7,6
    7 XA=FLOAT(K)/FLOAT(M)
      Z(I,2)=MNP*ABS(XA-YA)
      K=K+1
      GO TO 5
    6 YA=FLOAT(J)/FLOAT(N)
      Z(I,2)=MNP*ABS(XA-YA)
      J=J+1
    5 CONTINUE
      CMN=0.0
      I=1
    8 CMN=AMAX1(CMN,Z(I,2))
      I=I+1
      IF(I.LE.MN) GO TO 8
      CMN=CMN/FLOAT(MNP)
      RETURN
      END
      FUNCTION AKSCDF(M,N,D,U)
      DIMENSION U(N)
      K=FLOAT(M*N)*D+0.5
      U(1)=1.0
      DO 1 J=1,N
      U(J+1)=1.0
      IF(M*J.GT.K)U(J+1)=0.0
    1 CONTINUE
      DO 2 I=1,M
      W=FLOAT(I)/FLOAT(I+N)
      U(1)=W*U(1)
      IF(N*I.GT.K) U(1)=0.0
      DO 2 J=1,N
      U(J+1)=U(J)+U(J+1)*W
      IF(IABS(N*I-M*J).GT.K) U(J+1)=0.0
    2 CONTINUE
      AKSCDF=U(N+1)
      RETURN
      END
      SUBROUTINE SORT(A,N)
      DIMENSION A(1)
      KK=N-1
      DO 1 I=1,KK
      K=I+1
      DO 2 J=K,N
      IF(A(J).GT.A(I)) GO TO 2
      TEMP=A(J)
      A(J)=A(I)
      A(I)=TEMP
    2 CONTINUE
    1 CONTINUE
      RETURN
      END
```

1. This program computes the exact probability that the K-S criterion (D_{mn}) will be greater than or equal to the value of the observed test statistic (C/mn) [i.e., $P(D_{mn} \geq C/mn) = 1-p$]

2. (p) is the probability that an error will be committed if the hypothesis of equality of performance (H_o) is rejected. Therefore (1-p) is the confidence with which we reject H_o if we choose to do so.

Sequence 3

Variable	Wilk's Lamda	$F_{(1, 26)}$
Inventory	.9668	.8929
Mental Maturity Language	.9986	.0353
Mental Maturity Non Language	.9710	.7760
Personality Personal	.9936	.1669
Personality Social	.9925	.1957
Prior Exposure	.9711	2.5355

This table lists two statistics for each variable difference score. The first is Wilk's Lamda (λ), $0 \leq \lambda \leq 1$. As $\lambda \rightarrow 1$, the probability that the observed difference in means is likekly to have occurred by chance alone approaches one, in the limit. The second is the univariate F-ratio with appropriate degrees of freedom. For this statistic to be significant it must be greater than one. Note that only across "Prior Exposure" does this statistic approach statistical significance. If we were to reject H_o across the variable we would be forced to do so at some confidence level less than 90%. A level that low is normally considered insufficient for statistical/experimental work.

Sequence 4

Analysis of Variance (ANOV) to determine significance of observed differences in student performance across six-input variables by student by group.

Model

$$Y_{(ijk)} = \mu + G_i + e_{ik} + C_j + CG_{(ij)} + e_{(ijk)}$$

where: $Y_{(ijk)}$: the observed performance of the i^{th} group across the j^{th} variable by the k^{th} student.

μ : the grand mean

$(G_i + e_{(ik)})$: "Between-student" effects

where: G_i : Group $i = \begin{cases} 1 = \text{PSI} \\ 2 = \text{CLR} \end{cases}$

C_{ik} : Random error $k = \begin{cases} 1, \ldots, 12 & ; \text{if } i = 1 \\ 13, \ldots, 28 & ; \text{if } i = 2 \end{cases}$

$C_j + CG_{(ij)} + e_{(ijk)}$: within student effects

where: C_j = characteristic variable j : j=1, . . , 6

e_{ijk} = random error

All levels are fixed.

Legend:

Source:	the entity which could be generating observed variations
df :	degrees of freedom associated with the source
SS :	seem of squares of observed variates
MS :	mean square
EMS :	expected mean square
$F_{.90}(v1, v2)$:	the ratio that constituted the test statistic.

Sequence 4 (continued)

Source	df	SS	MS	EMS	$F_{(v_{i1} v_i)}$
Between Students$_R$	27	2864.33	106.09	.97	N.S.
Groups$_i$	1	10.71	10.71	.10	N.S.
e_{ir}	26	2853.52	109.75		
Within Students	140	136398.45	974.27	35.17	Sig.*
Characteristic Variables$_j$	5	132059.48	26411.90	953.50	Sig.
*$(C \times G)_{ij}$	5	738.33	147.67	5.33	Sig.
e_{ijk}	130	3600.64	27.70		
TOTAL	167	139262.78			

Source	df	SS	MS	EMS	$F_{(v_1 v_2)}$
Students	27	2864.33	106.09	.11	N.S.
Error	140	136398.45	974.27		

*Sig: significant at 90% level of confidence.

Sequence 5a

	Student	FE Raw Score		%
PSI Group	S_1	190/200	=	.95
	S_2	199	=	.99*
	S_3	173	=	.87*
	S_4	171	=	.86
	S_5	199	=	.99*
	S_6	188	=	.94
	S_7	192	=	.96
	S_8	194	=	.97*
	S_9	198	=	.99*
CLR Group	S_{10}	189/200	=	.95
	S_{11}	131	=	.66
	S_{12}	158	=	.79
	S_{13}	179	=	.90
	S_{14}	176	=	.88
	S_{15}	158	=	.79
	S_{16}	200	=	1.00
	S_{17}	142	=	.71
	S_{18}	170	=	.85
	S_{19}	180	=	.90
	S_{20}	159	=	.80

$\overline{X}_{PSI}$ = 189.33 = 95%

Std. dev = 10.58

*denotes students who took the FE early

$\overline{X}_{CLR}$ = 164.37 = 82%

Std. dev = 22.02

Sequence 5a (continued)

Student	Raw Score		%
S_{21}	173	=	.87
S_{22}	114	=	.57
S_{23}	151	=	.75
S_{24}	171	=	.86
S_{25}	179	=	.90

Statistical analysis of FE data

H_o: there is no difference in the performance of the two groups as measured by Final Exam Scores (FE).

H_1: otherwise

Confidence level = 90%

1. According to the non-parametric K-S test of the equivalence of two samples:

   Test Statistic (T.S.) = .653

   $P(D_{mn} \geq T.S.) = 1\text{-}p = .995 > .90$

Conclusion: reject H_o; the groups did not perform equivalently across F.E.

2. According to the parametric univariate F-ratio$_{(1,26)}$

   Test Statistic (T.S.) = 10.1017

   $P(F_{(1,26)} \geq T.S.) = 1\text{-}p = .996 > .90$

Conclusion: reject H_o.

Summary: the performance of the PSI group is significantly superior that of CLR. H_o could be rejected with more than 99% confidence wit[h] either test.

Sequence 5b

Course Grade
(C.G.)

Group	Student	C.G.		
PSI Group	S_1	4.00	;	(A)
	S_2	4.00	;	(A)
	S_3	3.66	;	(A-)
	S_4	3.66	;	(A-)
	S_5	4.00	;	(A)
	S_6	4.00	;	(A)
	S_7	4.00	;	(A)
	S_8	4.00	;	(A)
	S_9	4.00	;	(A)
CLR Group	S_{10}	4.00	;	(A)
	S_{11}	3.00	;	(B)
	S_{12}	3.33	;	(B+)
	S_{13}	4.00	;	(A)
	S_{14}	3.66	;	(A-)
	S_{15}	3.66	;	(A-)
	S_{16}	4.00	;	(A)
	S_{17}	3.00	;	(B)
	S_{18}	3.33	;	(B+)
	S_{19}	4.00	;	(A)
	S_{20}	3.33	;	(B+)

$\overline{X}_{PSI} = 3.92$; (A-)

Std. dev. = .150

$\overline{X}_{CLR} = 3.56$; (B+)

Std. dev. = .435

Sequence 5b (continued)

Student	C. G.		
S_{21}	4.00	;	(A)
S_{22}	2.66	;	(B-)
S_{23}	3.33	;	(B+)
S_{24}	3.66	;	(A-)
S_{25}	4.00	;	(A)

Statistical analysis of C. G. data

H_o: there is no difference in the performance of the two groups as measured by the Course Grade (C. G.)

H_1: otherwise

Confidence level = 90%

1. According to the non-parametric K-S test of the equivalence of two samples

   Test Statistic (T.S.) = .438

   $P(D_{mn} \geq T.S.) = 1-p = .854 < .90$

Conclusion: accept H_o; the groups did perform equivalently across C. G.

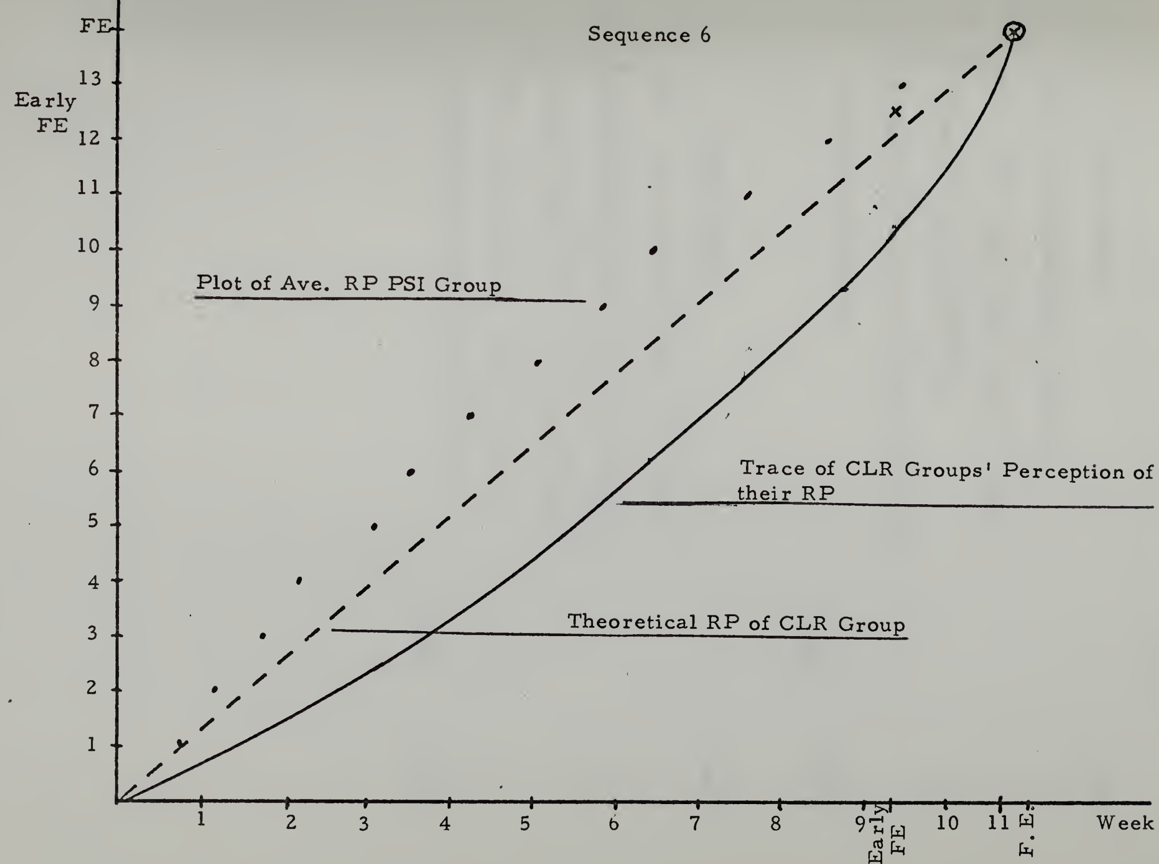
2. According to the parametric F-ratio $_{(1,26)}$

   Test Statistic (T.S.) = 5.8272

   $P(F_{(1,26)} \geq T.S.) = .975 > .90$

Conclusion: reject H_o.

Summary: non-parametric analysis denies rejection of H_o at 90% level. H_o could be rejected at the 85.4% level. Parametric analysis denies H_o at 97.5% confidence level.



Sequence 6
FE
13
Early FE
12
11
10
9
8
7
6
5
4
3
2
1
Plot of Ave. RP PSI Group
Trace of CLR Groups' Perception of their RP
Theoretical RP of CLR Group
1
2
3
4
5
6
7
8
9
Early FE
10
11
F. E.
Week

Sequence 7a

Typical Course Loads of Students By Group

PSI Group

Course Name		Credits
Intro. to Fortran Prog. (CS 2700)		3
Adv. Rev., Circ, Sigs, Sys. (EE 2106)		6
Digital Machines (EE 2810)		4.5
Computational Linear Algebra (MA 2045)		3
Differential Equations		4
	Total	20.5

CLR Group

Course Name		Credits
Elec. Engr. Fundamentals J (EE 2104)		4
Calculus and Vector Analysis (MA 1100)		6
Computational Linear Algebra (MA 2045)		3
Rev. Vector Mech., Fluids (PH 1051)		5
Programmable Calculators (ST 1810)		1
	Total	19.0

Sequence 7b

Group	GPA (w/o MA 2045)
PSI	3.76; (A-) (Std. Dev. = .223)
CLR	3.42; (B+) (Std. Dev. = .574)

Statistical Analysis

H_o : Group GPA's are equivalent

H_1 : otherwise

confidence level = .90

1 K-S Test:

T.S. = .500

$P(D_{mn} = T.S.) = .926 > .90$

2 Univariate $F\text{-ratio}_{(1,26)}$

G.S. = 3.358

$P(F_{(1,26)} = T.S.) = .925 > .90$

Summary: reject H_o

Sequence 8

<u>ANOV A</u> <u>L.A.</u> Students Output

<u>MODEL</u>

$$Y_{ijk} = \mu + G_i + e_{k(i)} + C_j + (CG)_{ij} + e_{k(ij)}$$

where μ = Grand Mean

$G_i + e_{k(i)}$ = effects between students

where : G_i = Group $\quad i = \begin{cases} 1 = \text{PSI} \\ 2 = \text{Convtl} \end{cases}$

$e_{k(i)}$ = Random Error $\quad k = \begin{cases} 1, - 9 & i = 1 \\ 10, - 25 & i = 2 \end{cases}$

$C_j + (CG)_{ij} + e_{k(ij)}$ = effects within students

where:

C_j = output variable $\quad j = 1, 2, 3$

$e_{k(ij)}$ = Random Error.

All levels fixed.

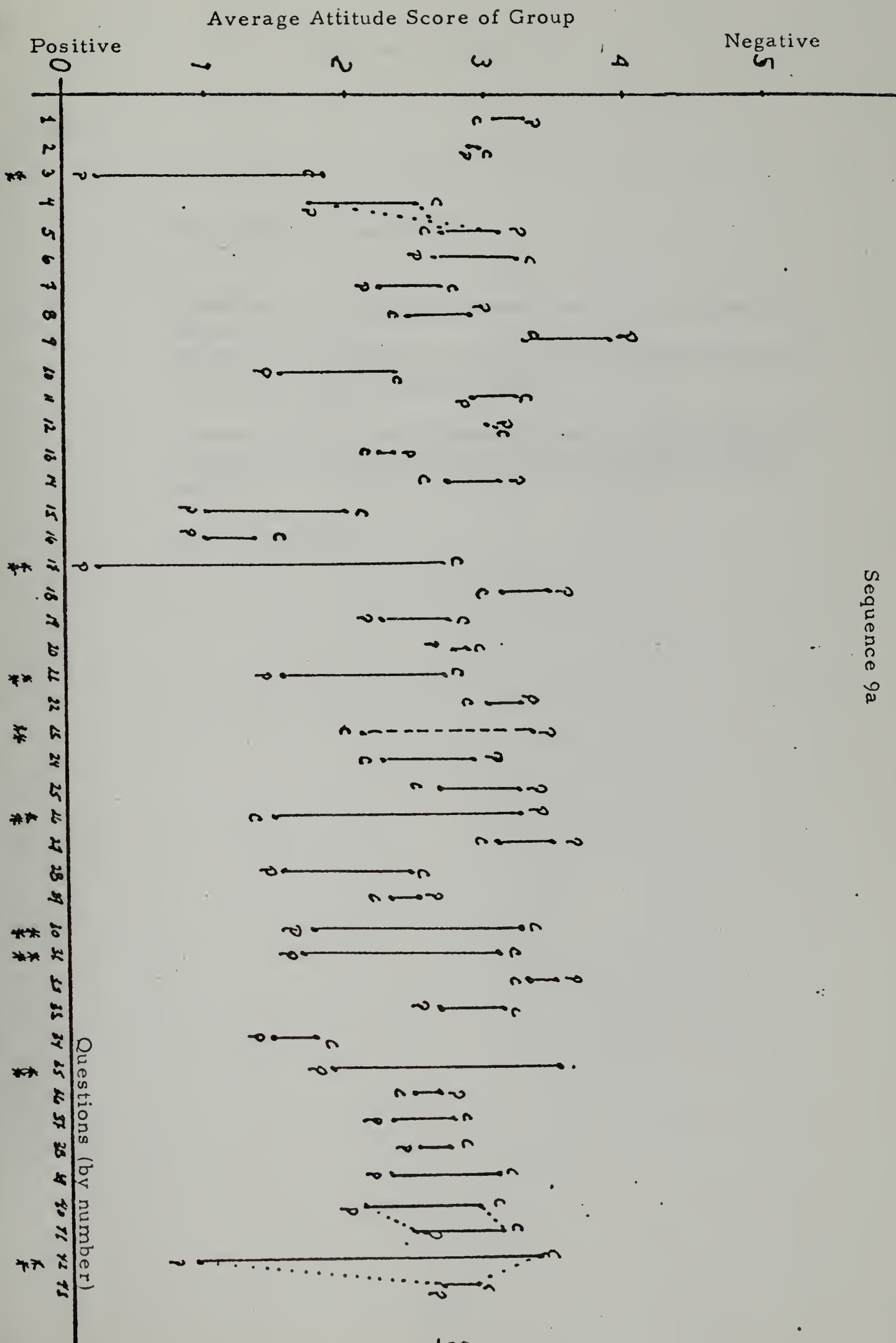
Table 10

Sequence 8 (continued)

Source	df	SS	MS	EMS	$F_{(v_1 v_2)}$	
Between	24	4171.05	173.79	1.39	N.S.	
Groups	1	1296.48	1296.48	10.37	Sig.	.01
$e_{k(i)}$	23	2874.57	124.98			
Within	50					
Output (var)	2	479021.92	239510.96	2047.78		
$(0 \times G)_{ij}$	2	2383.41	1191.71	10.19	Sig	= .01
$e_{k(ij)}$	46	5380.23	116.96			
Total	74	490956.61				

Source	df	SS	MS	EMS	$F(v_1 v_2)$
Students	24	4171.05	173.79	.02	N.S.
Error	50	48675.56	9735.71		
Total	74	490956.61			

Table 10 (continued)


Average Attitude Score of Group
Positive
Negative
0
1
2
3
4
5
Questions (by number)

Sequence 9a

Legend:

1. C: CLR Group Ave Response

   P: PSI Group Ave Response

2. *: Questions on which discriminant analysis of attitudes was based
   Significance of difference such that hypothesis of equality rejected at .90% confidence level.

   **: Question #23 where difference in favor of CLR group was statistically significant (also included in discriminant analysis).

3. ⁄ : Trend questions.

Sequence 9b

Attitude Survey Questionnaire

Question	Mean Score * PSI	CLR	Significance $F_{p=.10}(1,20)$
1. Memorization of basic facts in this course, is sufficient to obtain a "B" grade.	3.33	3.07	0.20
2. I would recommend that a friend take this course under the instructional technique I was exposed to (i.e., PSI or Conventional).	2.88	2.92	0.00
3. This course required that I assume active responsibility for my own learning.	0.22	1.84	14.57**
4. Prior to this course my study habits were	1.77	2.53	1.31
5. Right now my study efficiency is	3.11	2.76	0.28
6. I feel that the subject matter of this course was difficult to understand.	2.66	3.23	0.83
7. I am not satisfied with my understanding of the content of the course.	2.22	2.69	0.48
8. This course encouraged the development and use of my reasoning skill.	2.88	2.46	0.42
9. I was often frustrated in my efforts to understand the course material.	3.88	3.30	0.75
10. I have learned very little from this course.	1.55	2.38	1.77

		Mean Score PSI	Lecture	Significance $F_{p=.10}(1,20)$
11.	Taking this course has increased my interest in the subject area.	2.88	3.23	0.33
12.	The instructional technique employed in this course is an effective method for teaching the subject matter.	3.00	3.00	0.00
13.	I felt a keen sense of competition with my classmates.	2.22	2.30	0.01
14.	My main reason for working hard in this course was to obtain a grade of "B" or better.	3.11	2.76	0.31
15.	Too much material was covered in this course.	1.00	2.00	2.31
16.	I had difficulty getting assistance when I needed it.	1.00	1.38	0.82
17.	If I was asked for clarification of a point by a classmate, I felt very good about giving him assistance.	0.22	2.69	16.09**
18.	This course made me want to explore the subject area in greater depth.	3.44	3.15	0.18
19.	The instructional technique employed encouraged cooperation with my classmates.	2.22	2.76	1.14
20.	I found that taking this course was an unpleasant experience.	2.77	2.84	0.01
21.	I look forward to applying the techniques I learned on Practical Problems.	1.55	2.69	2.92**
22.	Most courses, in similar subjects, should be conducted as this one was.	3.22	3.00	0.14

		Mean Score		Significance
		PSI	Lecture	$F_{p=.10}(1,20)$
23.	The amount of knowledge I acquired was not worth the effort I expended.	3.33	2.15	3.10**
24.	In retrospect, taking this course was a worthwhile endeavor.	2.22	2.92	0.99
25.	I saw this course as an opportunity to learn rather than as a requirement to be met.	2.66	3.15	0.50
26.	The information presented in this course will be of value to the solution of practical problems.	1.44	3.23	11.11**
27.	Knowing what I do now, I would prefer to have taken this course via a different instructional technique.	3.44	3.07	0.23
28.	The conduct of this course accommodated my individuality (e.r., personal abilities, external distractions, conflicting commitments, etc.).	1.55	2.46	1.92
29.	The instructional technique did not foster my own desire to learn the material.	2.55	2.30	0.18
30.	The objective of this course were made clear.	1.77	3.23	5.36**
31.	Frequently I felt bored with this course.	1.66	3.07	3.59**
32.	My ability to grasp the subject matter exceeded my expectations.	3.55	3.30	0.16
33.	The objectives of this course were met.	2.66	3.15	0.85
34.	The course was of sufficient scope to challenge my academic abilities.	1.44	1.76	0.34

		Mean Score PSI	Lecture	Significance $F_{p=.10}(1,20)$
35.	Generally speaking, I felt I was pushed through the course at a rate faster than would allow me to absorb the material.	1.88	3.53	7.59**
36.	The primary motivating factor during this course was my own desire to acquire knowledge.	2.66	2.46	0.12
37.	I feel confident in my ability to do well on the final exam.	2.33	2.76	0.43
38.	This course, in its present form, could be taught effectively by any instructor qualified in the subject area.	2.55	2.69	0.04
39.	I believe that I have mastered the course material.	2.33	3.07	1.43
40.	When I began this course my prior attitudes towards tests were	2.11	2.92	1.53
41.	As the course progressed, my attitudes towards tests	2.44	3.07	1.28
42.	My prior motivation to learn this material (i.e., before the course even began) was	1.88	3.38	6.20**
43.	As the course progressed, my motivation seemed to	2.66	2.92	0.24

* A positive attitude is reflected by a lower score

** Indicates that the observed test statistic $F_{p=.10}(1,20)$ is significant at the .10 level of significance.

BIBLIOGRAPHY

1. Alba, E. and Pennypacker, H.S., "A Multiple Change Score Comparison of Traditional and Behavioral College Teaching Procedures," Journal of Applied Behavior Analysis, 1972, 5, 121-124.

2. Anderson, O.T., and Artman, R.A., "A Self-Paced, Independent Study, Introductory Physics Sequence -- Description and Evaluation," American Journal of Physics, 1972, 1737-1742.

3. Austin, S.M., and Gilbert, K.E., "Student Performance in a Keller Plan Course in Introductory Electricity and Magnetism," American Journal of Physics, 1973, 41, 12-18.

4. Austin, S.M., Gilbert, K.E., "Student Performance in a Keller Plan Course in Introductory Electricity and Magnetism," American Journal of Psychology (AJP), Vol. 11, 12-18 January 1973.

5. Bloom, B.S., "Learning for Mastery", UCLA Evaluation Comment 1968, 1 (Pt. 2).

6. Born, D.C., and Davis, M.L., "Amount and Distribution of Study in a Personalized Instruction Course and in a Lecture Course," Journal of Applied Behavior Analysis, 1972, 5 (1), 33-43.

7. Breland, N.S., and Smith, M.P., "A Comparison of PSI and Traditional Methods of Instruction for Teaching Introduction to Psychology," Personalized Instruction in Higher Education. Ruskin, R.S. and Bono, S.F. (Eds), Washington, D.C.: Center for Personalized Instruction, 1974.

8. Calhoun, J.F. "Elemental Analysis of the Keller Method of Instruction," State University of New York, Stony Brook. Paper presented at the annual meeting of the American Psychological Association, Montreal, August 1973.

9. Corey, J.R. and McMichael, J.S., "Retention in a PSI Introductory Psychology Course" PSI: 41 Germerical Papers, Sherman, J.G. (Ed), W. A. Benjamin Inc., Menlo Park, CA 1974.

10. Ferster, C.B., "Individualized Instruction in a Large Introductory Psychology College Course," Psychological Record, 1968, 18, 521-532.

11. Friedman, C.P., Hirschi, S., Parlett, M., Taylor, E.F., The Rise and Fall of PSI at MIT (unpublished, 1975).

12. Gallup, H.F., "The Introductory Psychology Course in Lafayette College, Easton, Pennsylvania: A Description and Tentative Evaluation." Paper presented at the annual meeting of the Eastern Psychological Association, 1970.

13. Gangne, R.M., "The Acquisition of Knowledge," Psychological Review, 1962, 69, 355-365.

14. Gangne, R.M., The Conditions of Learning, New York: Holt, Rinehart & Winston, 1965.

15. Green, B.A., "A Self-Paced Course in Freshman Physics," Occasional Paper #2, Education Research Center, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1969.

16. Green, B.A., "Is the Keller Plan Catching on Too Fast?" Journal of College Science Teaching, Vol. 1, No. 1, October 1971.

17. Johnson, J.M. and Pennypacker, H.S., "A Behavioral Approach to College Teaching," American Psychologist, 1971, 26, 219-244.

18. Keller, F.S., "A Personal Course in Psychology," In: The Control of Human Behavior, R. Ulrich, T. Stachnik and J. Mabry (Eds)., Glenview: Scott, Foresman, 1966.

19. Keller, F.S. and Sherman, J.G., The Keller Plan Handbook, Menlo Park, CA, W. A. Benjamin, Inc., 1974.

20. Kim, P.J., Jennrich, R.I., "The Exact Sampling Distribution of the Two-Sample Kolomogorov-Smirnov Criterion, D_{mn}; $m = n$" Selected Tables in Mathematical Statistics, Vol. 1, Harter, H. L., Owen, D.B., (Eds), Markham Publishing Company, Chicago, Illinois, 1970.

21. Kingsley, E.H., et al, "A Theoretical Basis for Individualized Instruction", Human Resources Research Organization AD-786 040 July 1974.

22. Klecka, W.R., "Discriminant Analysis," SPSS, 2d Ed., Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K., Bent, D.H., (Eds) McGraw Hill, 1975.

23. Koen, B.V., "The Keller Plan," Individualized Instruction in Engineering Education, Grayson, L.P., Beedenbach, J.M., (Eds) The American Society for Engineering Education, Washington, D.C., 1974.

24. Kulik, J.A., Kulik, C.L. and Milholland, J.E., "Evaluation of an Individualized Course in Psychological Statistics," Personalized Instruction in Higher Education, Reeskin, R.S., Bono, S.F. (Eds) Washington, D.C., Center for Personalized Instruction, 1974.

25. Kulik, J.A., "PSI: A Formative Evaluation" Invited Address at the National Conference on Personalized Instruction in Higher Education, Los Angeles, CA, March 1975.

26. Learmonth, G., "Computer Program to Comp the Exact Probabilities of the Occurrence of the Kolomgorov-Smirnov-Two-Sample, Non-Parametric, Criterion D_{mn}", unpublished, 1975 (NPS).

27. Lubkin, J.L., "Engineering Statics: A Keller-Plan Course with Novel Problems and Novel Features," Personalized Instruction in Higher Education, Reeskin, R.S., and Bono, S.F. (Eds), Washington, D.C.: Center for Personalized Instruction, 1974.

28. McMichael, J.S. and Corey, J.R., "Contingency Management in an Introductory Course Produces Better Learning," Journal of Applied Behavior Analysis, 1969, 2, 79-83.

29. McMichael, J.J., "Following-Up PSI Students," The Personalized System of Instruction Newsletter, Vol. 3, No. 1, Washington, D.C., Georgetown University, March 1975.

30. Moore, J.W., Hauck, W.E., Gangne, E.D., "Acquisition, Retention, and Transfer in an Individualized College Physics Course," Journal of Educational Psychology, Vol. 64, No. 3, 335-340, 1973.

31. Nazzaro, JR., Todorov, J.C., and Nazzaro, J.N., "Student Ability and Individualized Instruction," Journal of College Science Teaching, December 1972, 29-30.

32. Roth, C.H., Jr., "Continuing Effectiveness of Personalized Self-Paced Instruction in Digital Systems Engineering," Engineering Education, 1973, 63, (6).

33. Schimpfhauser, F., Horrocks, L., Richardson, K., Alben, J., Schumm, D., and Sprercher, "The Personalized Program of Instruction as an Adaptable Alternative Within the Traditional Structure of Medical Basic Sciences," Personalized Instruction in Higher Education, Washington, D.C., Center for Personalized Instruction, 1974.

34. Sheppard, W.D. and MacDermott, H.G., "Design and Evaluation of a Programmed Course in Introductory Psychology," Journal of Applied Behavior Analysis, 1970, 3, 511.

35. Silberman, R., Parker, Bruce, "Student Attitudes and The Keller Plan," Journal of Chemical Education, Vol. 51, No. 6, 393-395, June 1974.

36. Skinner, B.F., The Technology of Teaching, New York: Appleton-Century-Crofts, 1968.

37. Steinberg, D.I., Computational Matrix Algebra, McGraw Hill, New York, 1974.

38. Stice, J.E. (Ed), The Personalized System of Instruction: The Keller Plan in Engineering Education, (Bulletin No. 4), Austin, Texas, The University of Texas College of Engineering, 1971.

39. Sullivan, E.T., Clark, W.W., Tiegs, E.W., California Short-Form Test of Mental Maturity - College and Adult, 1963, S-Form, Level 5, California Test Board, McGraw Hill, Del Monte Research Park, Monterey, CA, 1963.

40. Thorpe, L.P., Clark, W.W., Teigs, E.W., California Test of Personality All levels, Grades: kgn to Adult, Form AA, California Test Board, McGraw-Hill, Del Monte Research Park, Monterey, CA, 1953.

41. Wagner, H., Behringer, R.P., "Individualized Course Comple-Tion Time Predictions: Development of Instruments and Techniques," HUMRO Technical Report 73-25, Nov 1973, HUMRO Division No. 1 (System Operations), Alexandria, Virginia.

42. Weissberg, H.L., "PSI - Innovation in the Engineering Classroom" Tennessee Engineering, 1973, 28-31.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0212 Naval Postgraduate School Monterey, CA 93940	2
3.	Department Chairman, Code 55So Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, CA 93940	1
4.	Professor Maurice D. Weir, Code 53Wc Department of Mathematics Naval Postgraduate School Monterey, CA 93940	3
5.	LT Gerald G. Brown, USN, Code 55Zr Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, CA 93940	1
6.	Honorable Wilbur B. Payne Deputy Under Secretary of the Army (OR) ODUSA(OR) SA WOOEAA The Pentagon Washington, DC 20310	1
7.	Dean of Educational Development/Executive Director Continuous Education Office (Code 024/500) Naval Postgraduate School Monterey, CA 93940	2
8.	Dr. Jack R. Borsting (02) Academic Dean Naval Postgraduate School Monterey, CA 93940	1
9.	RADM Isham Linder, USN (00) Superintendent Naval Postgraduate School Monterey, CA 93940	1
10.	Professor Ladis D. Kovach, Code 53Kv Chairman, Department of Mathematics Naval Postgraduate School Monterey, CA 93940	1

11. Dr. Ben A. Green — 1
Institute for Personalized Instruction
Georgetown University
Washington, DC

12. Professor Philip Doughty — 1
Institute of Technology
Syracuse University
Syracuse, New York

13. Dr. John Carter — 1
Naval Personnel Research Development Center
San Diego, CA

14. BGEN Alan R. Toffler, USA(Ret) — 1
V. P. National University
San Diego, CA

15. Captain Patrick A. Toffler, USA — 1
7453 Fairway Road
La Jolla, CA 92103

16. Chairman, Department of Meteorology (51) — 1
Naval Postgraduate School
Monterey, CA 93940

17. Chairman, Department of Electrical Engineering (52) — 1
Naval Postgraduate School
Monterey, CA 93940

18. Chairman, Department of Government (56) — 1
Naval Postgraduate School
Monterey, CA 93940

19. Chairman, Department of Aeronautics (57) — 1
Naval Postgraduate School
Monterey, CA 93940

20. Chairman, Department of Oceanography (58) — 1
Naval Postgraduate School
Monterey, CA 93940

21. Chairman, Department of Mechanical Engineering (59) — 1
Naval Postgraduate School
Monterey, CA 93940

22. Chairman, Department of Physics (61) — 1
Naval Postgraduate School
Monterey, CA 93940

23. Chairman, Operational Systems Technology (71) 1
Naval Postgraduate School
Monterey, CA 93940

24. Chairman, Computer Science Group (72) 1
Naval Postgraduate School
Monterey, CA 93940

25. Commanding General 1
U.S. Army Military Personnel Center
U.S. Army Civil Schools Branch
200 Stovall Street
Alexandria, VA 22332

162979

Thesis
T65
c.1

Toffler
An investigation of the relative effectiveness of the Personalized System of Instruction at the Naval Postgraduate School.